

KLASIFIKASI DATA MINING UNTUK PENERIMAAN SELEKSI CALON PEGAWAI NEGERI SIPIL 2014 MENGGUNAKAN ALGORITMA *DECISION TREE C4.5*

Rendragraha Kumara¹, Catur Supriyanto²

^{1,2}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang

Jl. Nakula I No. 5-11 Semarang Telp. (024) 3517261

E-mail : rendragrahakumara@gmail.com¹, catur.supriyanto@dsn.dinus.ac.id²

Abstrak

Badan Kepegawaian Daerah memegang wewenang dalam menjalankan visi dan misi daerah Kabupaten Banjarnegara salah satunya dengan pelaksanaan Rekrutmen dan penempatan yang sesuai syarat dan prosedur yang telah ditentukan oleh Pemerintah yang berwenang sesuai kebutuhan formasi yang dibentuk dalam setiap unit / organisasi badan. Banyak kasus dijumpai bahwa pemilihan pegawai dari tahun ke taun masih banyak yang tidak kompeten, tidak sesuai dengan kemampuan, dan bakat. Algoritma C4.5 digunakan untuk menentukan hasil lulus atau gagal nya seorang pelamar. Parameter pemilihan kelulusan adalah Indeks Prestasi Kumulatif, Umur, Akreditasi, Tes Wawasan Akademik, Tes Karakteristik Pribadi dan Tes Intelegensi Umum. Manfaat data mining dapat diimplementasikan pada penerimaan pegawai yang melibatkan data dalam jumlah besar, teknik yang digunakan dalam data mining ini adalah pohon keputusan, dengan Algoritma C4.5. Dari hasil penelitian yang menggunakan 250 data pelamar CPNS menghasilkan Hasil eksperimen dan evaluasi bahwa Algoritma Decision Tree C4.5 akurat diterapkan untuk penentuan diterimanya CPNS dengan tingkat akurasi terbaik yaitu 75 % dari tiga kali percobaan.

Kata Kunci: *Badan Kepegawaian Daerah, Kabupaten Banjarnegara, Rekrutmen dan Penempatan Pegawai, Klasifikasi, pohon keputusan. Algoritma C4.5*

Abstract

Regional Employment Board holds the authority to carry out the vision and mission areas Banjarnegara one of them with the implementation of appropriate recruitment and placement procedures and meet the requirements set by the competent authority in accordance formations formed in each unit / organizational entity. Many cases found that the selection of employees from year to epidemic are still many who are not competent, not according to their ability and talent. The use of data mining classification algorithm approach will be applied to determine how the suitability of the process with the results obtained, so the Regional Employment Board is not wrong to produce competent employees. C4.5 algorithms used to determine the results of pass or failure of an applicant. Parameter selection is graduation GPA, age, Accreditation, Academic Insights Test, Personal Characteristics Test and the General Intelligence Test. The benefits of data mining can be implemented on hiring that involve large amounts of data, the techniques used in data mining is a decision tree, the algorithm C4.5, after sorting the data, it will be seen how the tendency applicants will be accepted on recruitment, by searching The highest gain entropy and decision tree established. So in the implementation of selection recruitment, it is known patterns - patterns and criteria of what is needed in the implementation of recruitment and placement Governments Banjarnegara as a material classification. The experimental results and evaluations indicate that accurately C4.5 Decision Tree algorithm is applied to the determination of receipt CPNS with the best accuracy level that is 75% of the three attempts.

Keywords : *Regional Employment Board, Banjarnegara, Recruitment and Placement Officer, classifications, Decision tree. C4.5 Algorithm*

I. PENDAHULUAN

Badan Kepegawaian Daerah merupakan Badan Pemerintah yang berperan sebagai Penyelenggaraan Pengadaan dan Seleksi Calon Pegawai serta penempatan pegawai di lingkungan Pemerintah Kabupaten Banjarnegara. Salah satu tugas dan wewenang BKD atas pelaksanaan tugas kedinasan yang diberikan oleh Pemerintah Kabupaten Banjarnegara untuk melahirkan pegawai yang memiliki kemampuan baik dan kompeten di bidangnya yaitu dengan cara menyelenggarakan pelayanan umum bidang pengembangan pegawai, mutasi, pengangkatan dan kepangkatan, dalam rangka pelaksanaan urusan pemerintah bidang kepegawaian yang ditentukan dalam menjalankan tugas pengadaan dan manajemen kepegawaian sesuai formasi yang dibutuhkan setiap unit/organisasi badan.[7]

Melihat jumlah penduduk yang cukup padat (rasio beban kerja) dan pelamar luar daerah yang melamar, Pemerintah belum bisa memberi kepastian tentang jumlah Pegawai Negeri Sipil yang dibutuhkan dan yang harus direkrut kemudian diangkat setiap tahunnya masih belum dapat dihitung secara baik. Oleh sebab itu, alasan mengapa Pelaksanaan Pengadaan Pegawai, Mutasi dan Penempatan Pegawai di Lingkungan Pemerintah Kabupaten Banjarnegara diadakan. Salah satunya, berita dari tahun ke tahun dan opini yang beredar di kalangan masyarakat, penerimaan pegawai masih diwarnai kegiatan daerahisme, nepotisme, kolusi, spoil, sehingga kualitas SDM-PNS yang diterima sebagai pegawai masih rendah. [7]

Sebagai langkah dan solusi yang diajukan sebagai penyelesaian terhadap uraian diatas maka penulis tertarik untuk mengklasifikasi suatu permasalahan sistem data mining yang mudah digunakan dan dapat menyajikan informasi relevan dan akurat yaitu mengklasifikasi data mining untuk menampilkan informasi penerimaan

seleksi calon pegawai negeri sipil 2014 pada Badan Kepegawaian Daerah Kabupaten Banjarnegara menggunakan algoritma *Decision tree* C4.5.

Sehingga pelaksanaan pengadaan CPNS dapat berjalan sesuai prosedur yang ditetapkan oleh pemerintah. Sertadapat diketahui kriteria kriteria apa saja yang dibutuhkan dalam Pelaksanaan Pengadaan dan Seleksi Calon Pegawai Negeri Sipil di Lingkungan Pemerintah Kabupaten Banjarnegara sebagai bahan klasifikasi.

Data Mining sendiri memiliki beberapa algoritma, salah satunya yaitu *Decision Tree* yang merupakan metode klasifikasi yang paling terkenal, mudah dan banyak diimplementasi di berbagai bidang karena sederhana, serta memiliki kemampuan akurasi yang baik untuk mengklasifikasi gundukan data yang teramat banyak. Klasifikasi merupakan proses menemukan kumpulan pola atau fungsi-fungsi yang mendeskripsikan dan memisahkan kelas data satu dengan lainnya, untuk dapat digunakan untuk memprediksi data yang belum memiliki kelas data tertentu. Jadi secara singkat, klasifikasi adalah proses untuk membedakan/memisahkan kelas.

II. TEORI PENUNJANG

2.1 Business Intelligence

Business Intelligence (BI) is an umbrella term that combines architectures, tools, databases, analytical tools, applications, and methodologies [4]. BI meliputi semua proses mengumpulkan dan menganalisis data menggunakan teknologi yang bertujuan untuk mendapatkan informasi yang membantu sebuah institusi dalam mengambil keputusan [14], membantu organisasi mengelola dan menyaring informasi dalam membuat keputusan yang lebih efektif .[15]

BI menyediakan informasi yang bersifat historical, current dan predictive. Adapun fungsi-fungsi BI antara lain *reporting*,

online analytical processing, analytics, data mining, process mining, complex event processing, business performance management, benchmarking, text mining, predictive analytics dan prescriptive analytics.[17]

2.2 Data Mining

Data mining merupakan proses yang mempekerjakan satu atau lebih teknik pembelajaran *computer (machine learning)* untuk menganalisis dan mengekstraksi pengetahuan secara otomatis [18]. Data mining merupakan istilah yang sering dikatakan sebagai suatu cara untuk menguraikan serta mencari penemuan berupa pengetahuan didalam suatu *database*. Data mining adalah proses pemilihan atau “menambang” pengetahuan dari sekumpulan data dalam jumlah yang banyak.[2]

Pada Tabel 2.1 berikut mengklasifikasikan model data mining berdasarkan fungsi dan algoritma yang digunakan.

Tabel 2.1 : Data mining Functions, Algorithms, and Application Examples [4]

<i>Data mining function</i>	<i>Algorithm</i>	<i>Application Examples</i>
Association	Statistics, set theory	Market basket analysis
Classification	Decision trees, neural networks, control, risk assessment, rules	Target marketing quality
	Neural network, statistics, optimization, discriminate analysis	Market segmentation
Sequence discovery	Statistics, set theory	Market basket analysis over time, customer life cycle analysis
	Linear and nonlinear regression, curve fitting, neural networks	Sales forecasting, interest rate, prediction, inventory control
Drill-down and aggregate view of data	Visualization, using many different approaches.	Virtually all the preceding application

Berdasarkan penjelasan definisi tabel 2.1 , data mining dapat didefinisikan ke dalam faktor – faktor berikut :

1. Objek dari data mining adalah data kompleks atau yang berjumlah besar
2. Data mining adalah proses otomatis dari sekumpulan data yang terkumpul di masa lalu.
3. Tujuan data mining adalah menemukan pola atau hubungan hubungan yang bisa memberikan hasil indikasi yang bermanfaat.

2.3 Teknik Klasifikasi

Klasifikasi adalah salah satu bentuk dari teknik atau metode data mining yang termasuk dalam kategori *predictive mining* yaitu suatu teknik yang dapat digunakan untuk meramalkan atau memprediksi kecenderungan data di masa depan [3]. Proses yang terjadi dalam klasifikasi adalah proses penggolongan data ke dalam variabel target atau variabel tujuan dengan membangun sebuah model penyelesaian dengan memperhatikan atribut yang paling berpengaruh [5].

Klasifikasi masuk ke dalam *supervised induction*, dimana pengujian yang memanfaatkan kumpulan pengujian dari record dan atribut yang terklasifikasi untuk menentukan output dan kelas tambahan. Salah satu contoh algoritmanya adalah *decision tree* yang terkenal dan mudah dalam implementasinya ke dalam bentuk grafik.

Komponen-komponen utama dari proses klasifikasi antara lain :

- 1) Kelas, merupakan variable tidak bebas yang merupakan label dari hasil klasifikasi.
- 2) Prediktor, merupakan variable bebas suatu model berdasarkan dari karakteristik atribut data yang diklasifikasi,
- 3) Set data pelatihan, merupakan sekumpulan data lengkap yang berisi kelas dan predictor untuk dilatih agar model dapat mengelompokan ke dalam kelas yang tepat. Contohnya adalah grup pasien yang telah di-test terhadap serangan jantung, grup

pelanggan di suatu supermarket, dan sebagainya.

- 4) Set data uji, berisi data-data baru yang akan dikelompokkan oleh model guna mengetahui akurasi dari model yang telah dibuat.

2.4 Algoritma C4.5

Algoritma pohon keputusan yang terkenal adalah C4.5. Pada akhir tahun 1970 sampai awal tahun 1980 J. Ross Quinlan, seorang peneliti di bidang *machine learning*, membuat sebuah algoritma *decision tree* yang dikenal dengan ID3 (*Iterative Dichotomiser*). Quinlan kemudian mengembangkan algoritma ID3 menjadi algoritma C4.5 yang merupakan penyempurnaan algoritma sebelumnya [11].

Algoritma ini memiliki kelebihan, yaitu mudah dimengerti, fleksibel, dan menarik karena dapat divisualisasikan dalam bentuk gambar (pohon keputusan) [7]. Algoritma C4.5 membuat pohon keputusan dari node atas hingga cabang terakhirnya, dimana atribut paling atas merupakan akar, dan yang paling bawah dinamakan daun. Karena algoritma C4.5 digunakan untuk melakukan klasifikasi, jadi hasil daripengolahan test dataset berupa pengelompokkan data ke dalam kelas-kelasnya [14].

Rumus menghitung entropy pada algoritma C4.5

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

Keterangan :

- S : adalah Himpunan (dataset) kasus
- K : adalah banyaknya partisi S
- Pi : adalah probabilitaas yang didapat dari Sum(Ya) atau Sum(Tidak) dibagi total kasus

Setelah mendapatkan entropy dari keseluruhan kasus, lakukan analisis pada setiap atribut dan nilai nilainya dan kemudian hitung entropinya.

Kemudian hitung Gain dengan Rumus sebagai berikut :

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

S : himpunan kasus

|Si| : jumlah kasus pada partisi ke-i

A : atribut

|S| : jumlah kasus dalam S

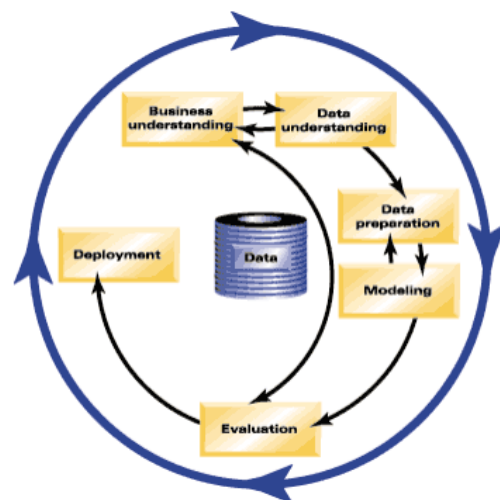
Langkah algoritma C4.5 untuk membangun pohon keputusan secara umum adalah sebagai berikut:

- a. Memilih atribut paling atas sebagai akar.
- b. Pilih cabang untuk setiap nilai.
- c. Tiap cabang bagi kasusnya .
- d. Ulangi proses pada setiap cabang hingga semua kasus yang ada pada cabang mempunyai kelas yang sama.

III. METODE PENELITIAN

3.1 Desain Penelitian

Standar proses data mining model CRISP-DM (*Cross Industry Standard Process*) yang dikembangkan tahun 1996 oleh analis dari beberapa industri menetapkan sebagai proses standar strategi pemecahan masalah data mining untuk penelitian. Dalam CRISP-DM, terdapat enam tahap siklus pengembangan data mining / tahapan proses sebagai berikut:



Gambar 3.1 : CRISP-DM [3]

a. Pemahaman Bisnis (*Business Understanding*)

Pemerintah melakukan rekrutmen CPNS merupakan kegiatan untuk mengisi lowongan yang kosong. Pada umumnya Pemerintah mengangkat hanya untuk menggantikan pegawai dipecat atau berhenti, yang sudah pensiun, serta meninggal dunia dengan tidak menambah jumlah pegawai secara keseluruhan yang kemudian ditetapkan dalam kebijakan nasional dalam keputusan Menteri di bidang pendayagunaan aparatur Negara.

b. Pemahaman Data (*Data Understanding*)

Data diperoleh dari Instansi Pemerintah Badan Kepegawaian Daerah Kabupaten Banjarnegara pada Bidang Jabatan oleh Kasubbid Jabatan sebagai sumber data primer sebanyak 250 sebagai bahan uji penelitian.

Atribut yang dijadikan sebagai acuan adalah indeks prestasi kumulatif, akreditasi kelulusan, hasil Tes Kompetensi Dasar (TKD) diantaranya : Tes Karakteristik Pribadi (TKP), Tes Wawasan Kebangsaan (TWK), Tes Intelegensi Umum (TIU) dan atribut lulus seleksi yaitu Diterima atau Gagal sebagai keterangan bahwa calon pegawai tersebut masuk sebagai pegawai atau tidak.

c. Pengolahan Data (*Data Preparation*)

Data kemudian dilakukan pemilihan atribut dan sebagian dari data dalam atribut yang ada akan dikonversikan untuk memudahkan proses data mining, karena data akan diproses dengan tools bantu data mining.

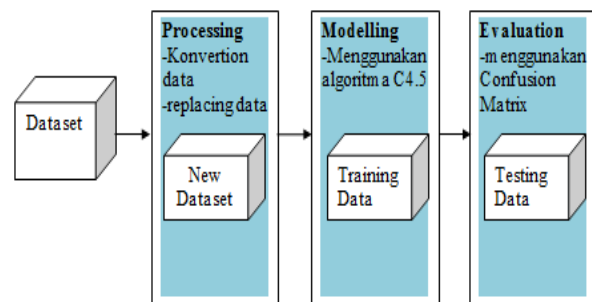
Tabel 3.1 Seleksi Atribut

Atribut	Detail Penggunaan	
NIK	×	No
No Register	×	No
No Peserta	√	Nilai Unique
Nama	×	No
Jenis Kelamin	×	No

Tempat Lahir	×	No
Tanggal Lahir	√	Nilai Model
Kode Pos	×	No
Propinsi	×	No
Kota	×	No
Telepon	×	No
Email	×	No
Asal Instansi	×	No
No Ijasah	×	No
Akreditasi	√	Nilai Model
Nilai IPK	√	Nilai Model
Kode Jabatan	×	No
Nama Jabatan	×	No
Kode Pendidikan	×	No
Nama Pendidikan	×	No
TWK	√	Nilai Model
TKP	√	Nilai Model
TIU	√	Nilai Model
Hasil Seleksi	√	Label Target

d. Pemodelan (*Modelling*)

Metode yang diusulkan dalam proses pemodelan adalah Algoritma C4.5. Kemudian pengukuran akurasi dalam penelitian ini akan menggunakan framework Matlab.



Gambar 3.2 Model penelitian yang diusulkan

e. Validasi dan Evaluasi

Dalam tahapan ini akan dilakukan validasi serta pengukuran keakuratan hasil yang dicapai oleh model menggunakan beberapa teknik yang terdapat dalam framework Matlab

R2010a yaitu Confusion Matrix untuk pengukuran tingkat akurasi model, dan pengujian manual.

f. Penyebaran (Deployment)

Hasil dari penelitian ini berupa analisa yang mengarah ke decision Support System (DSS), yang diharapkan dapat digunakan oleh Badan Kepegawaian Daerah Kab. Bnjarnegara Dalam menyeleksi dari ribuan pelamar Calon Pegawai Negeri Sipil yang akan mengisi kekosongan lowongan di Kab. Banjarnegara tersebut dan juga dapat digunakan sebagai bahan rujukan untuk penelitian selanjutnya.

IV. HASIL DAN PEMBAHASAN

Table 4.1: Data Pelamar CPNS 2014

No	No Peserta	Akreditasi	Umur	IPK	Nilai			Hasil
					TKP	TIU	TWK	
1	64183000023	B	23	3.09	144	75	75	Lulus
2	64183000033	B	23	3.46	149	110	70	Lulus
3	64183000040	A	28	3.19	155	105	85	Lulus
4	64183000042	B	26	3.08	150	70	70	Gagal
5	64183000048	B	22	3.6	149	95	80	Lulus
6	64183000052	B	27	3.04	145	65	95	Gagal
7	64183000054	B	24	3.24	147	85	90	Lulus
8	64183000055	B	26	3.26	149	95	75	Lulus
9	64183000057	B	31	3.36	147	80	85	Lulus
10	64183000060	B	32	3.32	139	105	80	Lulus
...
246	64183001199	A	24	3.13	146	80	65	Gagal
247	64183001211	B	22	3.05	148	90	50	Gagal
248	64183001216	A	28	3.34	143	100	75	Lulus
249	64183001217	B	28	3.11	148	125	100	Lulus
250	64183001218	A	25	3.34	145	95	60	Gagal

a. Konversi Data

Proses konversi data asli pelamar menjadi sebuah bilangan bulat untuk mempermudah matlab untuk melakukan proses klasifikasi.

Tabel 4.2 : Data pelamar setelah dikonversi

No	Akreditasi	Umur	IPK	Nilai			Hasil
				TKP	TIU	TWK	
1	2	1	2	1	2	2	Lulus
2	2	1	2	1	1	2	Lulus
3	1	2	2	1	1	2	Lulus
4	2	1	2	1	2	2	Gagal
5	2	1	1	1	2	2	Lulus
...
245	2	1	2	1	2	2	Gagal
246	1	1	2	1	2	2	Gagal
247	2	1	2	1	2	2	Gagal
248	1	2	2	1	1	2	Lulus
249	2	2	2	1	1	1	Lulus
250	1	1	2	1	2	2	Gagal

Konversi data ini merupakan proses penting dalam perhitungan dalam sistem yang dibangun agar memudahkan pengkodean dalam pembuatan sistem.

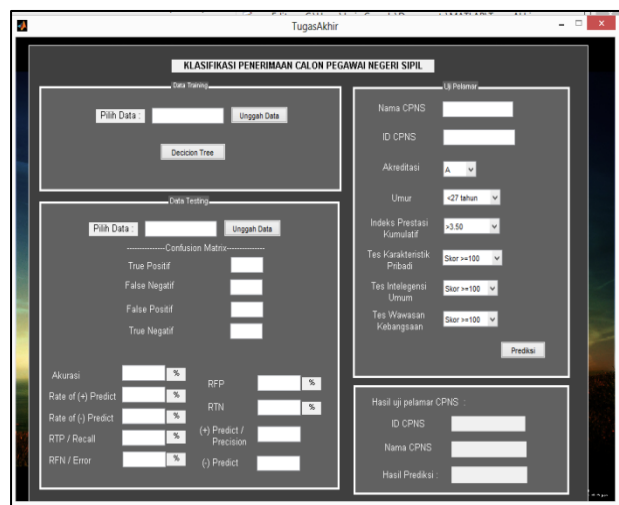
Dari Tabel 4.2 di atas diberikan penjelasan keterangan dari table tersebut diatas yaitu diambil dari enam variabel terpenting sebagai atribut untuk menentukan label / keluaran yaitu sebagai berikut:

Table 4.3: Keterangan table konversi data

Nama variable	Keterangan	Kelompok
Akreditasi	A	1
	B	2
	C	3
Umur	18-27 tahun	1
	28-35 tahun	2
IPK	3.0 – 3.50	1
	>3.50	2
TKP	Nilai ≥ 100	1
	Nilai < 100	2
TIU	Nilai ≥ 100	1
	Nilai < 100	2
TWK	Nilai ≥ 100	1
	Nilai < 100	2

b. Implementasi Antarmuka

Tampilan pada gambar 4.1 merupakan tampilan utama yang menampilkan menu-menu yang dapat diakses oleh user. Desain sistem yang digunakan yaitu menggunakan aplikasi Matlab versi 7.10.0 yang mendukung pengoperasian data mining.

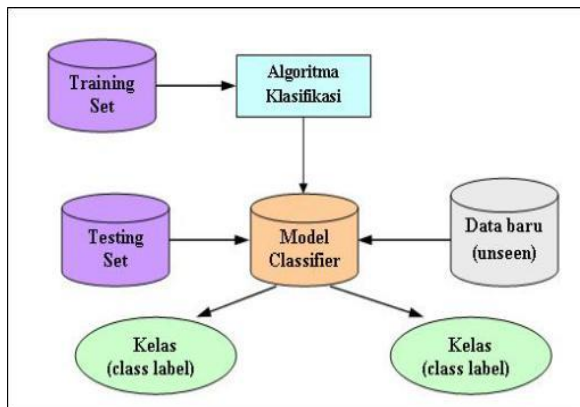


Gambar 4.1 : Tampilan Antarmuka Aplikasi Penerimaan CPNS

c. Proses Data Training dan Data Testing

Dalam proses data mining terdapat tahap yang harus dilakukan untuk pengolahan suatu data, yaitu tahap data training dan data testing. Training set digunakan oleh algoritma klasifikasi sebagai bentuk model classifier dalam bentuk pohon keputusan.

Testing set digunakan untuk mengukur classifier sejauh mana akurat melakukan klasifikasi dengan tepat. Algoritma C4.5 akan menghasilkan pola data yang diproses guna memberikan informasi dan trend dari data tersebut. Sedangkan data testing digunakan untuk mengukur sejauh mana pohon keputusan yang dibentuk berhasil melakukan klasifikasi dengan benar.



Gambar 4.2 : Proses Data

d. Pengujian Rules Terhadap Data Calon Siswa

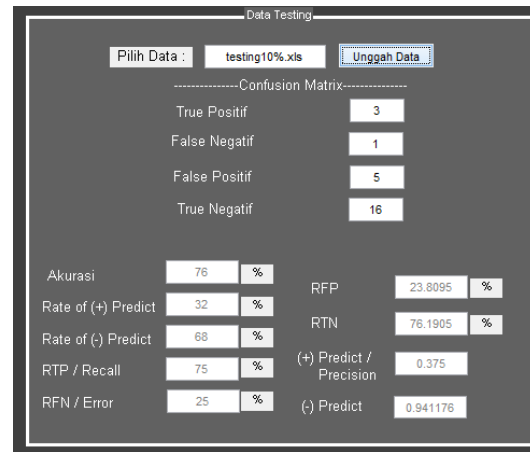
Pengujian terhadap validitas sistem bertujuan untuk mengetahui apakah solusi yang dihasilkan oleh pohon keputusan tersebut valid atau tidak. Rules tersebut dikatakan *valid* jika jumlah dan nama pelamar yang dinyatakan lulus BKD dengan *data set*.

Pembagian *data set* ke dalam dua bagian, yaitu *data training* dan *data testing*. Pengujian dilakukan tiga kali dengan jumlah *data training* dan *data testing* yang berbeda. Uraianya sebagai berikut :

1. *Data training* 90% dan *data testing* 10%

2. *Data training* 80% dan *data testing* 20%
3. *Data training* 70% dan *data testing* 30%

Pengujian I

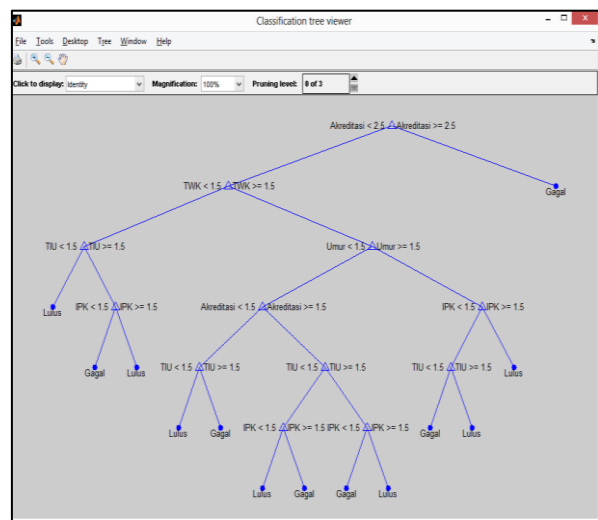


Gambar 4.3 : Hasil Data Testing 10%

Pada tabel 4.4 berikut, data dengan perbandingan 90% : 10% memiliki nilai akurasi sebesar 0.75, *precision* sebesar 0.375, *recall* sebesar 0.75 dan *f-measure* 0.5.

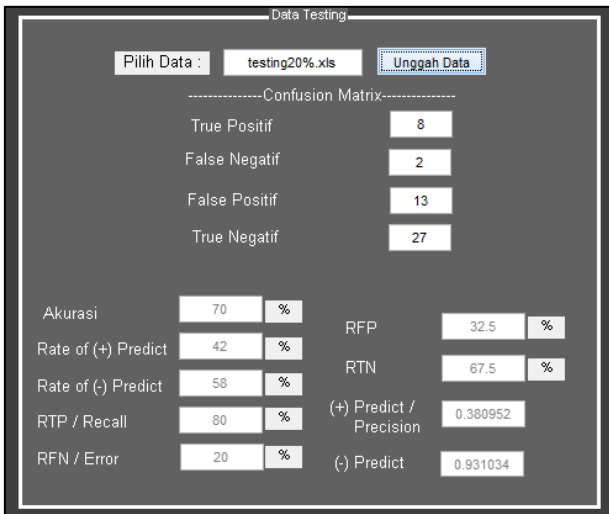
Tabel 4.4 Evaluasi dan Validasi 1

Persentase Data	Data training	Data testing	Akurasi	Precision	Recall	F-Measure
90% : 10%	225	25	0.75	0.375	0.75	0.5



Gambar 4.4 Pohon Keputusan uji I

Pengujian II



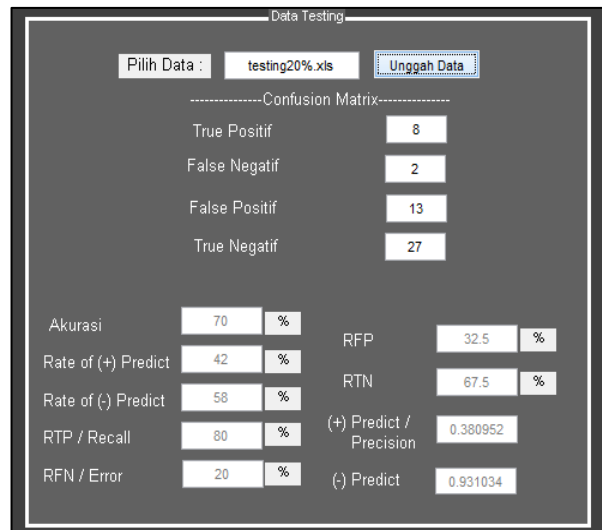
Gambar 4.5 : Hasil Data Testing 20%

Pada 4.5 berikut, data dengan perbandingan 80% : 20% memiliki nilai akurasi sebesar 0.75, *precision* sebesar 0.0.36, *recall* 0.8 dan *f-measure* 0.4965.

Tabel 4.5 Evaluasi dan Validasi II

Persentase Data	Data training	Data testing	Akurasi	Precision	Recall	F-Measure
80% : 20%	200	50	0.70	0.36	0.8	0.4965

Pengujian III

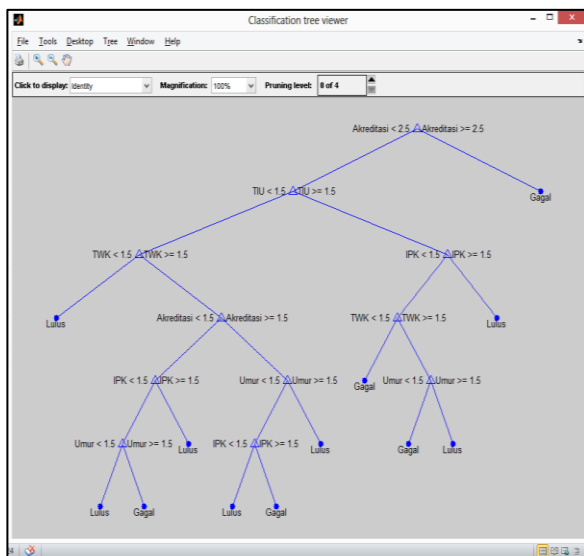


Gambar 4.7 : Hasil Data Testing 30%

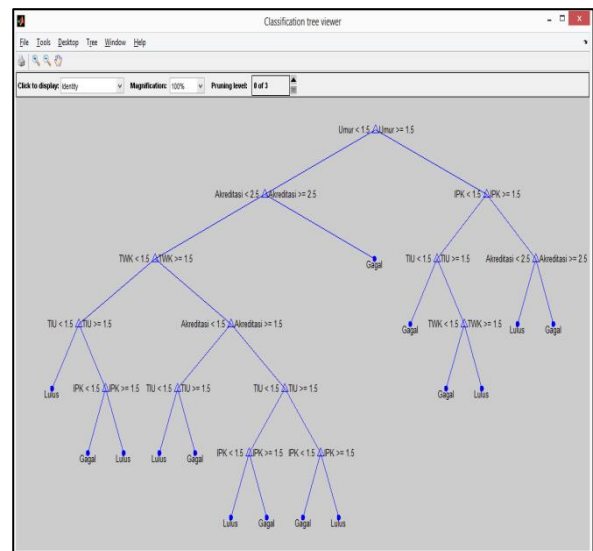
Pada 4.6 data dengan perbandingan 80% : 20% memiliki nilai akurasi sebesar 0.75, *precision* sebesar 0.0.36, *recall* 0.8 dan *f-measure* 0.4965.

Tabel 4.6 Evaluasi dan Validasi II

Persentase Data	Data training	Data testing	Akurasi	Precision	Recall	F-Measure
70% : 30%	175	75	0.6667	93.023	0.6	0.429

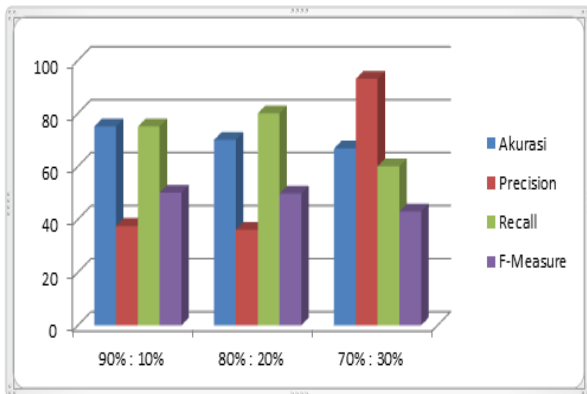


Gambar 4.6 Pohon Keputusan uji II



Gambar 4.8 Pohon Keputusan uji III

e. Analisa Hasil Pengujian



Gambar 4.9 Grafik Hasil Pengujian

Gambar 4.9 Grafik menunjukkan penurunan bertahap pada akurasi, dan *F-Measure*. *Recall* mengalami kenaikan dari 0.75 menjadi 0.8 namun turun menjadi 0.6.

Sedangkan *Precision* mengalami penurunan dari 0.375 menjadi 0.36 kemudian meningkat drastis menjadi 0.93. Semakin banyak *data training* yang digunakan maka semakin akurat hasil yang didapat.

V. KESIMPULAN DAN SARAN

a. Kesimpulan

Hasil penelitian yang didapat dari permasalahan yang dikembangkan disimpulkan bahwa pengimplementasian metode *decision tree* terhadap data penerimaan CPNS 2014 memiliki tingkat akurasi yang baik dalam menyelesaikan solusi klasifikasi dengan memanfaatkan teknik klasifikasi data mining dengan algoritma C4.5.

Hasil dari tiga kali pengujian, didapatkan akurasi terbaik dari pengujian pertama sebanyak 75% dari *dataset*. Sebanyak 250 pelamar yang dibagi menjadi 225 pelamar / 90% menjadi data training dan 25 pelamar / 10% menjadi data testing. Kemudian dibuatlah pohon keputusan dengan sistem yang dibangun, hubungannya untuk lebih memastikan dan mengetahui aturan atau *rules* yang

diperoleh dari perhitungan metode yang digunakan. Jadi dengan demikian metode *decision tree* merupakan metode yang sesuai dalam hal penyelesaian studi kasus penerimaan calon pegawai negeri sipil 2014.

b. Saran

Untuk pengembangan lebih lanjut maka penulis memberikan saran yang bermanfaat dan dapat membantu pembaca untuk masa yang akan datang, yaitu :

1. Menambahkan metode lain dalam data mining sebagai proses perbandingan tingkat keakuratan klasifikasi penerimaan CPNS 2014.
2. Penelitian selanjutnya hendaknya menggunakan data yang lebih banyak agar menghasilkan *rules* yang lebih akurat.
3. Pengujian metode ini belum sampai pada implementasi, maka perlu dibuat sistem aplikasi yang diperuntukkan untuk pihak Badan Kepegawaian Daerah Kab. Banjarnegara.

DAFTAR PUSTAKA

- [1] Mandasari, V., & Tama, B. A. *Analisis Kepuasan Konsumen Terhadap Restoran Cepat Saji Melalui Pendekatan Data Mining*, 2011.
- [2] Han, J. Kamber, *Data Mining: Concept and Techniques*. San Fransisco: Morgan Kaufmann Publishers, 2001.
- [3] Fayyad, Usama. *Advance in Knowledge Discovery and Data Mining*. MIT Press, 1996.
- [4] Turban, E., dkk. *Decision Support Systems and Intelligent Systems*. Yogyakarta: Andi Offset, 2005.
- [5] Kusriani. *Algoritma Data Mining*. Yogyakarta: Andi Offset, 2009.

- [6] Larose, Daniel T. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Willey & Sons, Inc, 2005.
- [7] Sugiat. *Peranan Badan Kepegawaian Daerah dalam pelaksanaan rekrutmen dan penempatan pegawai negeri sipil di lingkungan pemerintah Kota Samarinda*. 2013.
- [8] Friedman JH: *Data Mining and Statistics: whats the Connection?* Dept. of Statistics and Stanford Linear Accelerator Center. Stanford university. Stanford: California.
- [9] Zaky, M. *Data Mining and Analysis, Fundamental Concepts and Algorithms*. Iowa: University of Iowa, 2003.
- [10] K.Hastuti. *Analisis komparasi algoritma klasifikasi data mining untuk prediksi mahasiswa non aktif*. Semantik, 2012.
- [11] I. H. Witten, E. Frank, and M. A. Hall. *Data mining practical machine learning tools and technique*. burlington: morgan kaufmann publisher, 2011.
- [12] Fayyad, Usama. *Advances in Knowledge Discovery and Data Mining*. MIT Press, 1996.
- [13] Wenefrida Tulit Ina. *Klasifikasi Data Rekam Medis Berdasarkan Kode Penyakit Internasional Menggunakan Algoritma C4.5*. *Media Elektro*, vol. 1, pp. 105-110, 2013.
- [14] Connolly, Thomas and Begg, Carolyn. *Database Systems A Practical Approach to Design, Implementation, and Management Fifth Edition*. Boston: Pearson Education, 2010.
- [15] Lonnqvist & Pirttimaki. *The measurement of business intelligence*, 206.
- [16] Effendi, Onong Uchjana. *Ilmu, Teori, dan Filsafat Data Mining*. Bandung: Citra Aditya Bakti, 2001.
- [17] http://en.wikipedia.org/wiki/Business_intelegence
- [18] Santosa Budi. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*, Yogyakarta : Graha Ilmu. 2007.