



PENERAPAN MODEL *MACHINE LEARNING* UNTUK MENENTUKAN KLASIFIKASI JENIS BANTUAN SOSIAL

Nurvelly Rosanti¹, Muhammad Iqbal², Sirojul Munir³

^{1,2}Teknik Informatika, Universitas Muhammadiyah Jakarta

³Teknik Informatika, Sekolah Tinggi Teknologi Terpadu Nurul Fikri

Jakarta Pusat, DKI Jakarta, Indonesia 10510

Jakarta Selatan, DKI Jakarta, Indonesia 12640

nurvelly.rosanti@ftumj.ac.id, iqbalrohis@gmail.com, rojulman@nurulfikri.ac.id

Abstract

The Provincial Government of DKI Jakarta has a social assistance program budgeted by the APBD in the form of the Jakarta Elderly Card (KLJ), Jakarta Persons with Disabilities Card (KPDJ) and Jakarta Child Card (KAJ) programs. The problems that occur at the Kelurahan level are related to social assistance, namely the difficulty in determining the right type of assistance to be received by residents according to the terms and criteria that have been determined by the Government and there is no overlapping of recipients of assistance. The registration factor and the lack of understanding of residents regarding the criteria for the type of social assistance resulted in the determination of recipients of social assistance not being on target, such as residents receiving assistance who did not meet the criteria, resulting in social jealousy. To help with this problem, research was carried out to determine the best model in classifying the types of social assistance based on recipient criteria by comparing three classification methods. This study uses 100 respondent data and 8 criteria used as determinants of recipients. Comparison of the Certainty Factor, Naïve Bayes and Decision Tree models will provide an overview of the best model based on the level of accuracy. The confusion matrix is used to test the accuracy for Naïve Bayes and Decision Tree and the output of the selected model is a web-based application that can provide recommendations for types of social assistance. The best accuracy results are Certainty Factor which is 98.4%, Naïve Bayes and Decision Tree is 93.3%.

Keywords: Certainty Factor, Confusion Matrix, Decision Tree, Naïve Bayes, Social Assistance

Abstrak

Pemerintah Provinsi DKI Jakarta memiliki program bantuan sosial yang dianggarkan oleh APBD berupa program Kartu Lansia Jakarta (KLJ), Kartu Penyandang Disabilitas Jakarta (KPDJ) dan Kartu Anak Jakarta (KAJ). Permasalahan yang terjadi di tingkat Kelurahan terkait bantuan sosial yaitu kesulitan menentukan jenis bantuan yang tepat diperoleh warga sesuai dengan syarat dan kriteria yang sudah ditentukan Pemerintah serta tidak terjadi tumpang tindih penerima bantuan. Faktor pendaftaran serta kurangnya pemahaman warga terkait kriteria jenis bantuan sosial mengakibatkan penentuan penerima bantuan sosial belum tepat sasaran seperti warga penerima bantuan yang tidak sesuai dengan kriteria sehingga mengakibatkan kecemburuan sosial. Untuk membantu permasalahan tersebut dilakukan penelitian untuk menentukan model terbaik dalam klasifikasi jenis bantuan sosial berdasarkan kriteria penerima dengan membandingkan tiga metode klasifikasi. Penelitian ini menggunakan 100 data responden dan 8 kriteria yang digunakan sebagai penentu penerima. Perbandingan model *Certainty Factor*, *Naïve Bayes* dan *Decision Tree* akan memberikan gambaran model terbaik berdasarkan tingkat akurasi. *Confusion matrix* digunakan untuk menguji akurasi untuk *Naïve Bayes* dan *Decision Tree* dan luaran dari model yang terpilih adalah aplikasi berbasis web yang dapat memberikan rekomendasi jenis bantuan sosial. Hasil akurasi terbaik adalah *Certainty Factor* yaitu 98,4%, *Naïve Bayes* dan *Decision Tree* adalah 93,3%.

Kata kunci: Bantuan Sosial, *Certainty Factor*, *Confusion Matrix*, *Decision Tree*, *Naïve Bayes*

1. PENDAHULUAN

Kemiskinan merupakan fenomena sosial yang pasti terjadi di setiap negara. Secara umum, kemiskinan didefinisikan sebagai ketidakmampuan orang untuk memenuhi kebutuhan

ekonomi, sosial dan standar kebutuhan yang lain. Kemiskinan terus menjadi masalah fenomenal di belahan dunia, khususnya Indonesia yang merupakan Negara berkembang [1]. Salah satu permasalahan yang masih

dihadapi oleh bangsa Indonesia yaitu masalah kemiskinan. Pemerintah Indonesia telah banyak memiliki program-program untuk pengentasan kemiskinan. Beberapa aspek penting untuk mendukung keputusan strategi penanggulangan kemiskinan adalah tersedianya data kemiskinan yang akurat. Ketika data telah tersedia, maka pemerintah dapat mengambil keputusan apa saja yang harus dilakukan untuk penanggulangan tersebut. Selain itu, data yang tersedia dapat membuat pemerintah membandingkan angka kemiskinan dari tahun ke tahun. Bantuan Sosial merupakan pengeluaran berupa uang, barang, atau jasa yang diberikan oleh pemerintah pusat atau daerah kepada masyarakat guna melindungi masyarakat dari kemungkinan terjadinya risiko sosial, meningkatkan kemampuan ekonomi, serta kesejahteraan masyarakat [2].

Pemerintah Provinsi DKI Jakarta memiliki program bantuan sosial yang dianggarkan oleh APBD [3]. Program bantuan sosial Pemerintah Provinsi DKI Jakarta yaitu melalui pemberian Kartu Lansia Jakarta (KLJ), Kartu Penyandang Disabilitas Jakarta (KPDJ), dan Kartu Anak Jakarta (KAJ). Tujuan pemberian Kartu Lansia Jakarta (KLJ), Kartu Penyandang Disabilitas Jakarta (KPDJ), dan Kartu Anak Jakarta (KAJ) yaitu membantu penerima lansia untuk dapat memenuhi kebutuhan dasar serta meningkatkan kesejahteraan lanjut usia kemudian untuk penyandang disabilitas membantu mereka untuk dapat memenuhi kebutuhan dasar hidup dan mencegah anak dari risiko guncangan dan kerentanan sosial agar kelangsungan hidup terpenuhi.

Permasalahan di tingkat wilayah terkait dalam bantuan sosial terjadi disebabkan oleh beberapa faktor seperti pengambilan data belum akurat, penentuan jenis bantuan sosial belum tepat sasaran yang tidak sesuai kriteria yang mengakibatkan kecemburuan sosial dan kurangnya pemahaman warga terkait kriteria jenis bantuan sosial. Kriteria atau persyaratan untuk mendapatkan bantuan sudah ada dalam peraturan hanya kadang kala ada tumpang tindih bantuan seperti penerima bantuan lansia dan penyandang disabilitas.

Untuk mengatasi permasalahan tersebut, maka dilakukan penelitian untuk menentukan jenis bantuan sosial yang sesuai dengan kriteria dan tepat sasaran. Dengan menggunakan metode kecerdasan buatan untuk membantu mengklasifikasikan jenis bantuan sosial yang diperoleh warga berdasarkan syarat atau kriteria yang sudah ditentukan.

Penelitian sebelumnya terkait bantuan sosial oleh Damhuri (2021) yaitu klasifikasi kelayakan penerima bantuan sembako telah dikaji dengan menggunakan metode *naïve bayes* [4]. Pada penelitian tersebut tentang kelayakan penerima sembako dengan kriteria keluarga miskin dengan dilihat kondisi penghasilan, kepemilikan rumah, jumlah tanggungan dengan hasil pengujian akurasi model 86%. Penelitian lain terkait bantuan sosial juga sudah dibahas oleh Qadrini (2021) yaitu klasifikasi penerima bantuan

sosial menggunakan *Decision Tree* dan *Adaboost*, kasus ini lebih detail tentang penerima bantuan dampak Covid-19. Kriteria penerima berupa kondisi rumah tangga yaitu penghasilan, pekerjaan, sumber air dan kondisi rumah, dengan akurasi model adalah *Decision Tree* 94% dan 95% untuk *Adaboost* [5].

Banyak metode dalam mengimplementasikan kecerdasan buatan ke dalam sebuah sistem sesuai dengan bentuk data yang diperoleh. Banyak metode klasifikasi yang dapat digunakan seperti pada penelitian ini menggunakan *Decision Tree*, *Naïve Bayes* dan *Certainty Factor*. Penerapan klasifikasi juga dapat menggunakan sistem pakar dengan menggunakan metode *Certainty Factor*, hal ini terdapat pada penelitian Jamaludin (2013) yaitu tentang klasifikasi anak kebutuhan khusus [6].

Sehubungan sudah adanya penelitian terkait bantuan sosial menggunakan *Naïve Bayes* dan *Decision Tree* walaupun kriteria dan sasaran klasifikasi berbeda maka metode ini juga dapat diterapkan pada penerima bantuan sosial di Pemerintahan DKI Jakarta khusus bantuan sosial Kartu Lansia Jakarta (KLJ), Kartu Penyandang Disabilitas Jakarta (KPDJ), dan Kartu Anak Jakarta (KAJ). Di samping kedua metode tersebut juga dapat dilakukan penelitian perbandingan terhadap penerapan sistem pakar untuk kasus ini dengan menggunakan metode *Certainty Factor*. Pada *Certainty Factor* untuk menentukan jenis atau golongan bantuan sosial khususnya adalah Menentukan jenis bantuan sosial, menentukan kriteria dari jenis bantuan sosial, menentukan *rule*, melakukan perhitungan pada *Certainty Factor*, dan memperoleh hasil.

Untuk memperoleh metode yang terbaik dalam menentukan klasifikasi penentuan penerima bantuan sosial KLJ, KPDJ dan KAJ dilakukan membandingkan tiga model yaitu model sistem pakar menggunakan *Certainty Factor* dan menggunakan *machine learning* menggunakan metode *Naïve Bayes* dan *Decision Tree*. Berdasarkan hal tersebut di atas, maka dapat dilakukan menentukan model yang terbaik dan dapat membantu Satuan Pelaksana Sosial Kecamatan Cilincing Jakarta Utara, DKI Jakarta dalam menentukan penerima bantuan sosial secara sistem.

2. METODE PENELITIAN

Untuk menyelesaikan permasalahan ini maka pada penelitian ini melakukan langkah-langkah penelitian mulai dari pengumpulan data, pra-proses data, pengolahan data, pembuatan model, pengujian dan pembuatan aplikasi.

2.1 Metode Pengumpulan Data, Instrumen Penelitian, dan Metode Pengujian

Data jenis bantuan sosial ini diperoleh melalui wawancara dengan Kepala Satuan Pelaksana Sosial Kecamatan Cilincing. Wawancara dilakukan bertujuan mengetahui permasalahan terkait kesulitan dalam menentukan penerima dan distribusi bantuan sosial di daerah tersebut sehingga

dari hasil wawancara diperoleh data peserta penerima KLJ, KPDJ dan KAJ.

Data peserta penerima KLJ, KPDJ dan KAJ yang diperoleh berbentuk tabular. Data yang digunakan sebanyak 100 data penerima KLJ, KPDJ dan KAJ. Data yang diperoleh tersebut sudah dapat ditentukan fitur-fitur sebagai syarat atau kriteria penerima yang digunakan memiliki tipe data numerik dan nominal seperti pada Tabel 1 dan jenis bantuan sosial terdapat pada Tabel 2

Tabel 1. Kriteria dan Tipe Data Penerima Bantuan Sosial

| No | Fitur | Jenis Data |
|----|-------------------------------------|------------|
| 1. | Umur | Numerik |
| 2. | DTKS | Nominal |
| 3. | Hasil Verifikasi Pendamsos | Nominal |
| 4. | Penyanggah Disabilitas atau tidak | Nominal |
| 5. | BUMN/PNS/TNI/POLRI/Anggota DPR/DPRD | Nominal |
| 6. | Rumah Tangga Memiliki Mobil | Nominal |
| 7. | Tanah/Lahan dengan NJOP 1 Miliar | Nominal |
| 8. | Rata-rata Penghasilan | Numerik |

Tabel 2. Jenis Bantuan Sosial

| No. | Jenis Bantuan Sosial |
|-----|---|
| 1. | Kartu Lansia Jakarta (KLJ) |
| 2. | Kartu Penyandang Disabilitas Jakarta (KPDJ) |
| 3. | Kartu Anak Jakarta (KAJ) |

Pada Tabel 3 terdapat syarat atau kriteria penerima bantuan sosial baik KLJ, KPDJ dan KAJ. Untuk KLJ dan KAJ prioritas penentu adalah fitur umur.

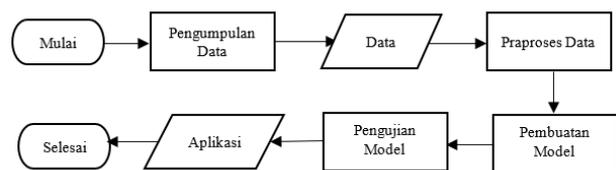
Tabel 3. Kriteria Penerima Bantuan Sosial

| No | Jenis Bantuan Sosial | Kriteria |
|----|----------------------------|---|
| 1. | Kartu Lansia Jakarta (KLJ) | <ul style="list-style-type: none"> - Warga lansia yang berumur ≥ 60 tahun - Dalam Status Ekonomi terendah dan terdaftar dalam Data Terpadu Kesejahteraan Sosial (DTKS). - Hasil verifikasi dari pendamping sosial (pendamsos) Kelurahan. - Tidak memiliki mobil - Tidak ada anggota PNS/TNI/POLRI/DPR/DPRD - Tidak memiliki tanah/lahan (dengan NJOP 1 Miliar) - Penghasilan \leq Rp 2.000.000. |

| No | Jenis Bantuan Sosial | Kriteria |
|----|---|---|
| 2. | Kartu Penyandang Disabilitas Jakarta (KPDJ) | <ul style="list-style-type: none"> - Dalam Status Ekonomiterendah dan terdaftar dalam Data Terpadu Kesejahteraan Sosial (DTKS). - Penyandang disabilitas dari keluarga prasejahtera. - Hasil verifikasi dari pendamping sosial (pendamsos) Kelurahan - Tidak memiliki mobil - Tidak ada anggota PNS/TNI/POLRI/DPR/DPRD - Tidak memiliki tanah/lahan (dengan NJOP 1 Miliar) - Penghasilan \leq Rp 2.000.000. |
| 3. | Kartu Anak Jakarta (KAJ) | <ul style="list-style-type: none"> - Anak usia dini berusia 0-6 tahun - Dalam Status Ekonomi terendah dan terdaftar dalam Data Terpadu Kesejahteraan Sosial (DTKS) - Hasil verifikasi dari pendamping sosial (pendamsos) Kelurahan - Tidak memiliki mobil - Tidak ada anggota PNS/TNI/POLRI/DPR/DPRD - Tidak memiliki tanah/lahan (dengan NJOP 1 Miliar) - Penghasilan orang tua \leq Rp 2.000.000. |

2.2 Tahapan Penelitian

Tahapan dalam penelitian ini dimulai dari pengumpulan data, menelaah data, pra-proses data, pembuatan model, pengujian model dan pembuatan aplikasi, seperti pada Gambar 1.



Gambar 1. Tahapan Penelitian

Pada Gambar 1 dijelaskan tahapan-tahapan penelitian yang terdiri dari :

- a. Pengumpulan data
Tahap ini dilakukan wawancara langsung kepada Kepala Satuan Pelaksana Sosial Kecamatan Cilincing Jakarta Utara. Hasil wawancara diperoleh informasi terkait 100 responden penerima bantuan sosial.
- b. Data
Data yang diperoleh sudah berupa data tabular dengan fitur-fitur atau variabel seperti pada Tabel 1 dan Tabel 2.
- c. Pra-proses Data

Setelah memperoleh data maka tahap berikutnya menelaah data berupa pengecekan terhadap fitur-fitur apakah terdapat data *missing value*, tipe data fitur dan pengecekan distribusi sebaran data agar tidak

mempengaruhi performa model. Dari telaah data tidak ditemukan data *missing*, semua baris terisi data kemudian sebaran data juga ditemukan distribusi normal seperti pada Gambar 2. [7]

| Name | Type | Missing | Statistics | | Filter (9 / 9 attributes): Search for Attributes |
|-------------------------------------|-------------|---------|-------------------------------|------------------------------|--|
| UMUR | Integer | 0 | Min 0 | Max 79 | Average 29.450 |
| DTKS | Polynominal | 0 | Least TIDAK TERDAFTAR (50) | Most TERDAFTAR (50) | Values TERDAFTAR (50), TIDAK TERDAFTAR (50) |
| HASIL VERIFIKASI PENDAMSOS | Polynominal | 0 | Least TIDAK DITEMUKAN (50) | Most DITEMUKAN (50) | Values DITEMUKAN (50), TIDAK DITEMUKAN (50) |
| PENYANDANG DISABILITAS ATAU TIDAK | Polynominal | 0 | Least PENYANDA [...] TAS (27) | Most TIDAK PE [...] TAS (73) | Values TIDAK PE [...] ABILITAS (73), PENYANDANG |
| JENIS BANTUAN | Polynominal | 0 | Least KLJ (30) | Most KAJ (35) | Values KAJ (35), KP DJ (35), ...[1 more] |
| BUMN/PNS/TNI/POLRI/ANGGOTA DPR/DPRD | Polynominal | 0 | Least YA (19) | Most TIDAK (81) | Values TIDAK (81), YA (19) |
| RUMAH TANGGA MEMILIKI MOBIL | Polynominal | 0 | Least YA (11) | Most TIDAK (89) | Values TIDAK (89), YA (11) |

Gambar 2. Telaah Data

d. Pembuatan Model

Model pertama menggunakan sistem pakar yaitu menggunakan *Certainty Factor* (CF), model kedua menggunakan klasifikasi *Naïve Bayes* dan model ketiga menggunakan *Decision Tree*.

1. *Certainty Factor* (CF)

Untuk CF dibutuhkan pembobotan oleh pakar terkait penentuan tingkat keyakinan variabel atau fitur tersebut sebagai penentu dalam penentuan penerima bantuan sosial. Pembobotan dilakukan oleh Kepala Satuan Pelaksana Sosial Kecamatan Cilincing Jakarta Utara dengan hasil seperti pada Tabel 4. [8]

Tabel 4. Bobot CF

| No | Jenis Bantuan Sosial | Kriteria | Bobot |
|----|----------------------------|---|-------|
| 1. | Kartu Lansia Jakarta (KLJ) | - Warga lansia yang berumur ≥ 60 tahun | 0,8 |
| | | - Dalam Status Ekonomi terendah dan terdaftar dalam Data Terpadu Kesejahteraan Sosial (DTKS). | 0,6 |
| | | - Hasil verifikasi dari pendamping sosial (pemdamsos) Kelurahan | 0,4 |
| | | - Tidak memiliki mobil | 0,6 |
| | | - Ada anggota PNS/TNI/POLRI/DPR/DPRD | 0,2 |
| | | - Tidak memiliki tanah/lahan (dengan NJOP 1 Miliar) | 0,6 |
| | | - Penghasilan \leq Rp 2.000.000 | 0,8 |

| No | Jenis Bantuan Sosial | Kriteria | Bobot |
|--|---|---|--------------------------|
| 2. | Kartu Penyandang Disabilitas Jakarta (KPDJ) | - Dalam Status Ekonomi terendah dan terdaftar dalam Data Terpadu Kesejahteraan Sosial (DTKS). | 0,6 |
| | | - Penyandang disabilitas dari keluarga prasejahtera. | 1,0 |
| | | - Hasil verifikasi dari pendamping sosial (pendamsos) Kelurahan | 0,6 |
| | | - Tidak memiliki mobil | 0,6 |
| | | - Anggota PNS/TNI/POLRI/DPR/DPRD | 0,2 |
| | | - Tidak memiliki tanah/lahan (dengan NJOP 1 Miliar) | 0,6 |
| | | - Penghasilan \leq Rp 2.000.000. | 0,8 |
| | | 3. | Kartu Anak Jakarta (KAJ) |
| - Dalam Status Ekonomi terendah dan terdaftar dalam Data Terpadu Kesejahteraan Sosial (DTKS) | 0,6 | | |
| - Hasil verifikasi dari pendamping sosial (pendamsos) Kelurahan | 0,4 | | |
| - Tidak memiliki mobil | 0,6 | | |
| - Anggota PNS/TNI/POLRI/DPR/DPRD | 0,2 | | |
| - Tidak memiliki tanah/lahan (dengan NJOP 1 Miliar) | 0,6 | | |
| - Penghasilan orangtua \leq Rp 2.000.000 | 0,8 | | |

Kemudian langkah berikutnya pemberian kode fitur untuk penentuan *rules* seperti pada Tabel 5.

Tabel 5. Kode Fitur

| Kode | Fitur |
|------|---|
| A1 | Warga yang berumur ≥60 Tahun |
| A2 | Terdaftar dalam Data Terpadu Kesejahteraan |
| A3 | Hasil verifikasi dari pendamping sosial (pendamsos) Kelurahan |
| A4 | Tidak Memiliki mobil |
| A5 | Anggota PNS/TNI/POLRI/DPR/DPRD |
| A6 | Tidak memiliki tanah/lahan (dengan NJOP 1 Miliar) |
| A7 | Berpenghasilan ≤2.000.000. |
| A8 | Penyandang Disabilitas |
| A9 | Anak yang berusia 0-6 Tahun |

Rule 1 : IF A1 AND A2 AND A3 AND A4 AND A5 AND A6 AND A7 THEN B1

Rule 2 : IF A2 AND A3 AND A4 AND A5 AND A6 AND A7 AND A8 THEN B2

Rule 3 : IF A2 AND A3 AND A4 AND A5 AND A6 AND A7 AND A9 THEN B3

Sehingga derajat ketidakyakinan terhadap eviden diperoleh sebagai berikut,

$$CF(H)KLJ = \text{IF } A1 \{CF = 0.8\} \text{ AND } A2 \{CF = 0.6\} \text{ AND } A3 \{CF = 0.4\} \text{ AND } A4 \{CF = 0.6\} \text{ AND } A5 \{CF = 0.2\} \text{ AND } A6 \{CF = 0.6\} \text{ AND } A7 \{CF = 0.8\} \text{ THEN } B1 \{CF = 0.75\}$$

$$CF(H)KLJ = \min\{0.8, 0.6, 0.4, 0.6, 0.2, 0.6, 0.8\} \times 0.75$$

$$CF(H)KLJ = 0.2 \times 0.75$$

$$CF(H)KLJ = 0.15$$

$$CF(H)KPDJ = \text{IF } A2 \{CF = 0.6\} \text{ AND } A3 \{CF = 0.4\} \text{ AND } A4 \{CF = 0.6\} \text{ AND } A5 \{CF = 0.2\} \text{ AND } A6 \{CF = 0.6\} \text{ AND } A7 \{CF = 0.8\} \text{ AND } A8 \{CF = 1.0\} \text{ THEN } B1 \{CF = 0.70\}$$

$$CF(H)KPDJ = \min\{0.6, 0.4, 0.6, 0.2, 0.6, 0.8, 1.0\} \times 0.70$$

$$CF(H)KPDJ = 0.14$$

$$CF(H)KAJ = \text{IF } A9 \{CF = 0.4\} \text{ AND } A2 \{CF = 0.6\} \text{ AND } A3 \{CF = 0.4\} \text{ AND } A4 \{CF = 0.6\} \text{ AND } A5 \{CF = 0.2\} \text{ AND } A6 \{CF = 0.6\} \text{ AND } A7 \{CF = 0.8\} \text{ THEN } B1 \{CF = 0.65\}$$

$$CF(H)KAJ = \min\{0.4, 0.6, 0.4, 0.6, 0.2, 0.6, 0.8\} \times 0.65$$

$$CF(H)KAJ = 0.13$$

Sehingga diperoleh nilai ketidakyakinan terhadap fakta sangat kecil sekali yaitu untuk KLJ = 0,15, KPDJ = 0,14 dan KAJ = 0,13.

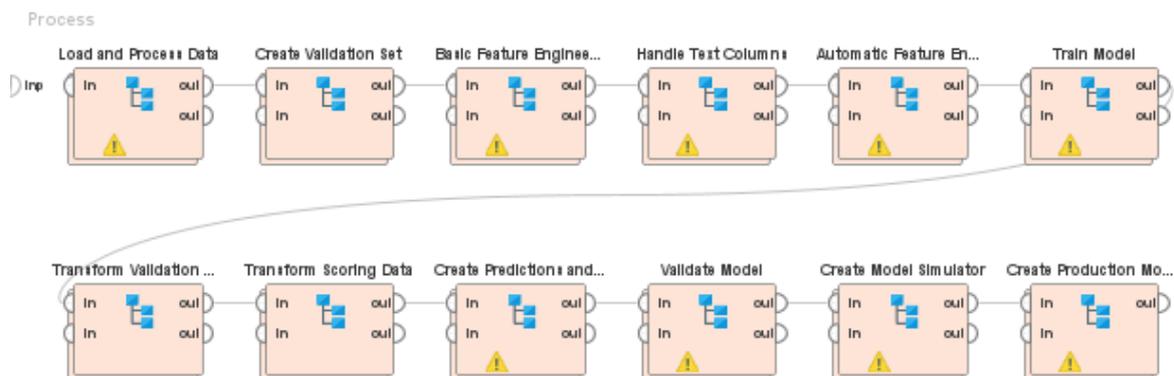
2. Naïve Bayes

Perancangan model klasifikasi dengan Naïve Bayes menggunakan perangkat *Rapid Miner* dengan memanfaatkan menu *Auto Model* sehingga mengikuti langkah-langkah seperti pada Gambar 3.



Gambar 3. Tahapan Auto Model untuk Naïve Bayes

Pada Gambar 3 dimulai dengan *Load Data* yaitu *input dataset* bantuan sosial kemudian dilakukan *Select Task* untuk menyeleksi atribut yang tidak digunakan yaitu atribut No dan Nama. Pada tahap *Prepare Target* dilakukan penentuan label kelas atau fitur target yaitu fitur Jenis Bantuan Sosial dengan kelas terdiri dari kelas KLJ, KPDJ dan KAJ. Tahap berikutnya *Select Inputs* untuk memilih fitur-fitur latih yang memiliki korelasi terhadap fitur target. Setelah persiapan pra-proses tersebut maka tahap berikutnya memilih metode klasifikasi yaitu memilih metode *Naïve Bayes*, sehingga diperoleh tahap-tahapan model pembuatan *Naïve Bayes* dapat dilihat pada Gambar 4. [9][10]



Gambar 4. Proses Pembuatan Model Naïve Bayes

Pada Gambar 4 dalam proses pembuatan model *Naïve Bayes* secara *Auto Model* maka dibutuhkan 12 tahapan dimulai dari mengambil *dataset* kemudian dilakukan validasi *dataset* sebelum dilakukan perhitungan *Naïve Bayes*. Pada validasi *dataset* ini terdapat proses *split* data untuk membagi data latih dan data uji yaitu membagi 60% data latih dan 40% data uji.

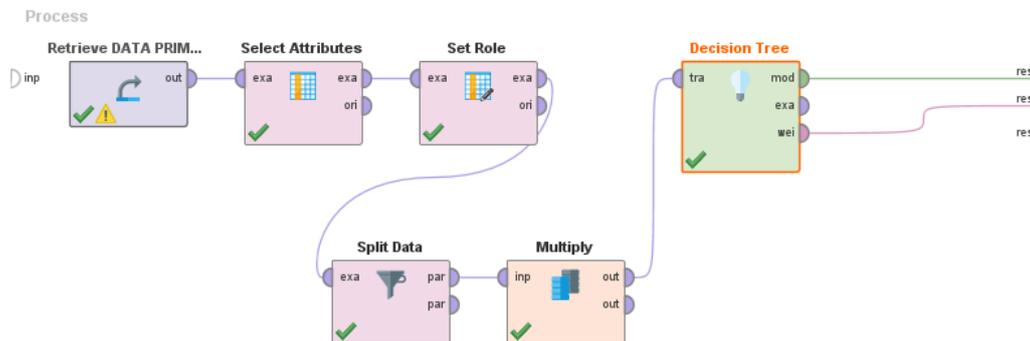
3. *Decision Tree*

Perancangan model dengan *Decision Tree* hampir sama dengan langkah *Naïve Bayes* pada Gambar 3 hanya pada

tahap *Model Types* memilih metode *Decision Tree*. Setelah menentukan fitur target yaitu fitur Jenis Bantuan Sosial dengan *output* klasifikasi KLJ, KPDJ dan KAJ [11]. Kemudian dihitung pembobotan fitur-fitur latih dengan metode *information gain* untuk penentuan *root tree* dengan hasil pembobotan seperti pada Gambar 5 dengan fitur Umur menjadi fitur penentu dalam model yaitu dengan bobot 0,207. Algoritma yang digunakan adalah C4.5, kemudian berikutnya dilakukan pembagian data menjadi data latih dan data uji dengan pembagian 80% data latih dan 20% data uji. [12][13]

| Attribute | Weight |
|--|--------|
| UMUR | 0.207 |
| TANAH/LAHAN (DENGAN NJOP DI ATAS 1 MILIAR) | 0.039 |
| DTKS | 0.038 |
| HASIL VERIFIKASI PENDAMSOS | 0.034 |
| BUMN/PNS/TNI/POLRI/ANGGOTA DPR/DPRD | 0.024 |
| PENYANDANG DISABILITAS ATAU TIDAK | 0.020 |
| RUMAH TANGGA MEMILIKI MOBIL | 0.011 |
| RATA-RATA PENGHASILAN | 0.004 |

Gambar 5. Pembobotan *Decision Tree*



Gambar 6. Tahapan Perancangan Model *Decision Tree*

Perancangan model menggunakan metode *Decision Tree* menggunakan perangkat *Rapid Miner* menu *Auto Model* dapat dilihat langkah-langkahnya pada Gambar 6.

Pada Gambar 6 dimulai dengan *input dataset* yang diperoleh dari Kelurahan yang terdiri dari 11 kolom dengan fitur yaitu no, nama lengkap, umur, memiliki mobil, rata-rata penghasilan, penyandang disabilitas, jenis bantuan, hasil verifikasi pendamsos, DTKS, PNS, tanah/lahan. Dataset ini terdiri dari 100 responden yang mengisi sesuai kriteria pertanyaan.

Tahap berikutnya adalah seleksi atribut yaitu untuk menghilangkan fitur yang tidak digunakan atau tidak berhubungan dengan pembentukan model. Fitur yang dihilangkan adalah fitur no dan nama karena sifatnya unik dan tidak bisa diklasifikasikan.

Setelah seleksi atribut atau fitur maka tahap berikutnya *set role* untuk pelabelan dengan menentukan fitur target klasifikasi yaitu terpilih fitur Jenis Bantuan yang terdapat 3 kelas yaitu KLJ, KPDJ dan KAJ.

Tahap berikutnya adalah *split data* untuk pembagian data set sebagai data latih dan data uji. Pembagian data menggunakan komposisi 80% untuk data latih dan 20% untuk data uji. Kemudian tahap akhir dari pembuatan model adalah menggunakan metode *decision tree* dengan parameter *information gain*.

e. Pengujian Model

Setelah pembuatan model maka dilakukan pengujian terhadap model baik terhadap CF, *Naïve Bayes* dan *Decision Tree*.

1. *Certainty Factor*

Pada CF akan diuji tingkat keyakinan (E) menggunakan input dari pengguna sebagai contoh :

Pengguna meng-input kriteria A1 AND A6 AND A7 Kemudian dilakukan pengujian menggunakan formula
 a. *Input yang sama dengan kriteria KLJ*

$$CF_c(CF1, CF6) = CF1 + CF6(1 - CF1)$$

$$CF_c(CF1, CF6) = 0,8 + 0,6(1 - 0,8)$$

$$CF_c(CF1, CF6) = 0,8 + 0,12$$

$$CF_c(CF1, CF6) = 0,92$$

$$CF_t(CF_c, CF7) = CF_c + CF7(1 - CF_c)$$

$$CF_c(CF1, CF7) = 0,92 + 0,8(1 - 0,92)$$

$$CF_c(CF1, CF7) = 0,92 + 0,064$$

$$CF_c(CF1, CF7) = 0,984$$

b. *Input yang sama dengan kriteria KPDJ*

$$CF_c(CF1, CF6) = CF1 + CF6(1 - CF1)$$

$$CF_c(CF1, CF6) = 0,0 + 0,6(1 - 0,0)$$

$$CF_c(CF1, CF6) = 0,0 + 0,6$$

$$CF_c(CF1, CF6) = 0,6$$

$$CF_t(CF_c, CF7) = CF_c + CF7(1 - CF_c)$$

$$CF_c(CF1, CF7) = 0,6 + 0,8(1 - 0,6)$$

$$CF_c(CF1, CF7) = 0,6 + 0,32$$

$$CF_c(CF1, CF7) = 0,92$$

c. *Input yang sama dengan kriteria KAJ*

$$CF_c(CF1, CF6) = CF1 + CF6(1 - CF1)$$

$$CF_c(CF1, CF6) = 0,0 + 0,6(1 - 0,0)$$

$$CF_c(CF1, CF6) = 0,0 + 0,6$$

$$CF_c(CF1, CF6) = 0,6$$

$$CF_t(CF_c, CF7) = CF_c + CF7(1 - CF_c)$$

$$CF_c(CF1, CF7) = 0,6 + 0,8(1 - 0,6)$$

$$CF_c(CF1, CF7) = 0,6 + 0,32$$

$$CF_c(CF1, CF7) = 0,92$$

Sehingga diperoleh tingkat keyakinan atau akurasi berdasarkan inputan pengguna bahwa 0,984 atau 98,4% adalah penerima KLJ.

2. *Naïve Bayes*

Pengujian performa model pada Naïve Bayes menggunakan *confusion matrix* dengan hasil pada Gambar 7 yaitu tingkat presisi KLJ dan KAJ sangat baik yaitu 100% artinya bahwa model dengan input oleh pengguna dapat tepat menentukan klasifikasi jenis bantuannya. Untuk KPDJ hanya memiliki presisi 85,71% dengan makna bahwa masih terdapat salah prediksi kelas dari *input data* yang diuji. Untuk *recall* kelas KLJ dan KPDJ bernilai 100% artinya bahwa data uji tepat sesuai kelas yang diprediksi oleh model dan untuk kelas KAJ *recall* 77,78% artinya dari 9 data hanya 7 yang tepat sesuai prediksi model. Hasil uji performa akurasi model secara keseluruhan adalah 93,3% seperti pada Gambar 8. [14]

| | true KAJ | true KLJ | true KPDJ | class precision |
|--------------|----------|----------|-----------|-----------------|
| pred. KAJ | 7 | 0 | 0 | 100.00% |
| pred. KLJ | 0 | 8 | 0 | 100.00% |
| pred. KPDJ | 2 | 0 | 12 | 85.71% |
| class recall | 77.78% | 100.00% | 100.00% | |

Gambar 7. Pengujian Performa Model Naïve Bayes

| Criterion | Value | Standard Deviation |
|----------------------|-------|--------------------|
| Accuracy | 93.3% | ± 9.1% |
| Classification Error | 6.7% | ± 9.1% |

Gambar 8. Hasil Akurasi Performa Model Naïve Bayes

3. *Decision Tree*

Pengujian performa model *Decision Tree* menggunakan *confusion matrix* dengan hasil performa seperti pada Gambar 9 yaitu sama dengan hasil pada *Naïve Bayes* dengan

tingkat presisi KLJ dan KAJ sangat baik yaitu 100% artinya bahwa model dengan *input* oleh pengguna dapat tepat menentukan klasifikasi jenis bantuannya. Untuk KPDJ hanya memiliki presisi 85,71%. Untuk *recall* kelas KLJ dan KPDJ bernilai 100% artinya bahwa data uji tepat sesuai

kelas yang diprediksi oleh model dan untuk kelas KAJ *recall* 77,78% artinya dari 9 data hanya 7 yang tepat sesuai

prediksi model. Hasil uji performa akurasi model secara keseluruhan adalah 93,3% seperti pada Gambar 10.

| | true KAJ | true KLJ | true KPDJ | class precision |
|--------------|----------|----------|-----------|-----------------|
| pred. KAJ | 7 | 0 | 0 | 100.00% |
| pred. KLJ | 0 | 8 | 0 | 100.00% |
| pred. KPDJ | 2 | 0 | 12 | 85.71% |
| class recall | 77.78% | 100.00% | 100.00% | |

Gambar 9. Pengujian Performa Model *Decision Tree*

| Criterion | Value | Standard Deviation |
|----------------------|-------|--------------------|
| Accuracy | 93.3% | ± 9.1% |
| Classification Error | 6.7% | ± 9.1% |

Gambar 10. Hasil Akurasi Performa Model *Decision Tree*



Gambar 11. Tampilan Aplikasi Berbasis Web

f. Pembuatan aplikasi
Setelah membuat model tahap berikutnya adalah *deployment* dengan merancang luaran dari model berupa aplikasi berbasis web dengan pemrograman PHP. Hasil tampilan dari aplikasi seperti pada Gambar 11. [15]

3. HASIL DAN PEMBAHASAN

Perbandingan dari tiga model yaitu *Certainty Factor*, *Naïve Bayes* dan *Decision* maka diperoleh performa akurasi dari masing-masing model seperti Tabel 6.

Tabel 6. Perbandingan Hasil Performa Akurasi

| | <i>Certainty Factor</i> | <i>Naïve Bayes</i> | <i>Decision Tree</i> |
|---------|-------------------------|--------------------|----------------------|
| Akurasi | 98,4% | 93,3% | 93,3% |

Pada Tabel 6 dapat dilihat bahwa hasil akurasi performa model yang paling baik adalah *Certainty Factor* yaitu 98,4%. Hanya saja model CF bergantung kepada intervensi pakar dalam menentukan bobot fitur apa yang paling penting dalam kriteria yang ada. Berdasarkan Tabel 3 maka

dapat dilihat bahwa memang faktor utama pembeda dari penerima bantuan sosial adalah umur untuk KLJ yaitu umur lebih besar dan sama dengan 60 tahun, sedangkan untuk KAJ umur 0-6 tahun. Untuk KPDJ menjadi syarat utama adalah penyandang disabilitas. Sehingga berdasarkan fitur penentu ini mengakibatkan tingkat akurasi dari model CF adalah 98,4%.

Untuk model *Naïve Bayes* dan *Decision Tree* menghasilkan performa model yang sama yaitu 93,3% sehingga model ini juga baik untuk penentuan klasifikasi tanpa intervensi pakar, cukup dari data yang sudah divalidasi oleh yang berwenang.

4. KESIMPULAN

Berdasarkan tujuan bahwa mencari model terbaik menggunakan metode *machine learning* untuk menentukan penerima bantuan sosial yaitu KLJ, KPDJ dan KAJ berdasarkan kriteria yang sudah ditentukan oleh pemerintah. Diperoleh hasil performa akurasi yang baik adalah sistem pakar dengan metode CF yaitu 98,4%.

Jika tidak ingin melibatkan pakar maka dapat menggunakan model *Naïve Bayes*, walaupun tingkat akurasi sama dengan *Decision Tree* tapi waktu proses lebih cepat untuk *Naïve Bayes* yaitu 427 ms sedangkan waktu proses *Decision Tree* 605 ms.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Kepala Satuan Pelaksana Sosial Kecamatan Cilincing Jakarta Utara, DKI Jakarta.

DAFTAR PUSTAKA

- [1] Y. P. Gratia, P. Issak Benyamin, Y. Sumarno, dan V. Wariki, "Pengembangan Model Pendidikan Agama Kristen Bagi Anak Korban Kemiskinan," *J. Ecodunamika*, vol. 3, no. 1, hal. 1, 2020.
- [2] D. Ardiansyah, W. Suharso, dan G. I. Marthasari, "Analisis Penerima Bantuan Sosial menggunakan Bayesian Belief Network," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 2, hal. 506–513, 2018, doi: 10.29207/resti.v2i2.447.
- [3] P. D. Pergub DKI Jakarta, "Gubernur Daerah Khusus Ibukota Jakarta," vol. 2019, no. 93, hal. 9–11, 2021.
- [4] A. Damuri, U. Riyanto, H. Rusdianto, dan M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *J. Ris. Komput.*, vol. 8, no. 6, hal. 219–225, 2021, doi: 10.30865/jurikom.v8i6.3655.
- [5] L. Qadrini, A. Seppewali, dan A. Aina, "Decision Tree dan Adaboost Pada Klasifikasi Penerima Program Bantuan Sosial," *J. Inov. Penelit.*, vol. 2, no. 7, hal. 1959–1966, 2021.
- [6] H. Jamaludin, "Aplikasi Metode Certainty Factor Pada Pengembangan Sistem Pengklasifikasi Anak Berkebutuhan Khusus," vol. 03, hal. 132–143, 2013.
- [7] J. P. Jiawei Han, Micheline Kamber, *Data Mining : Concepts and Technicques*, 3rd Edition. Morgan Kaufmann, 2012.
- [8] A. Sucipto, Y. Fernando, R. I. Borman, dan N. Mahmuda, "Penerapan Metode Certainty Factor Pada Diagnosa Penyakit Saraf Tulang Belakang," *J. Ilm. FIFO*, vol. 10, no. 2, hal. 18, 2019, doi: 10.22441/fifo.2018.v10i2.002.
- [9] D. Novianti, "Implementasi Algoritma Naïve Bayes Pada Data Set Hepatitis Menggunakan Rapid Miner," *Paradig. - J. Komput. dan Inform.*, vol. 21, no. 1, hal. 49–54, 2019, doi: 10.31294/p.v21i1.4979.
- [10] D. Wilandini dan Purwnatoro, "Penerapan Algoritma Naive Bayes Dalam Mengklasifikasikan Media Sosial Untuk Mengamati Trend Kuliner," vol. 8, no. 1, hal. 31–39, 2022, [Daring]. Tersedia pada: <https://journal.nurulfikri.ac.id/index.php/jtt>.
- [11] F. J. Kaunang, "Penerapan Algoritma J48 Decision Tree Untuk Analisis Tingkat Kemiskinan di Indonesia," *CogITO Smart J.*, vol. 4, no. 2, hal. 348, 2019, doi: 10.31154/cogito.v4i2.141.348-357.
- [12] V. Podgorelec dan M. Zorman, "Decision Tree Learning," in *Encyclopedia of Complexity and Systems Science*, 2015.
- [13] R. K. Amin, Indwiarti, dan Y. Sibaroni, "Implementasi Klasifikasi Decision Tree Dengan Algoritma C4 . 5 Dalam Pengambilan Keputusan Permohonan Kredit Oleh Debitur," *e-Proceeding Eng.*, vol. 2, no. 1, 2015.
- [14] S. Visa, B. Ramsay, A. Ralescu, dan E. Van der Knaap, "Edited by Sofia Visa, Atsushi Inoue, and Anca Ralescu," in *Maics*, 2011, vol. 710, hal. 120–127.
- [15] M. T. Prihandoyo, "Unified Modeling Language (UML) Model Untuk Pengembangan Sistem Informasi Akademik Berbasis Web," *J. Inform. J. Pengemb. IT*, vol. 3, no. 1, hal. 126–129, 2018.