

## Klasifikasi Kelayakan Keluarga Penerima Bantuan Langsung Tunai Menggunakan *Gain Ratio* Dan *Naïve Bayes*

Shelby Amalia Sandi<sup>1</sup>, Kurniabudi<sup>2</sup>, Yudi Novianto<sup>3</sup>

<sup>1,2,3</sup>Ilmu Komputer, Teknik Informatika, Universitas Dinamika Bangsa, Jambi, Indonesia

Email: <sup>1</sup>[shelby.amsan@gmail.com](mailto:shelby.amsan@gmail.com), <sup>2</sup>[kbudiz@yahoo.com](mailto:kbudiz@yahoo.com), <sup>3</sup>[Viant.yudi@gmail.com](mailto:Viant.yudi@gmail.com)

Email Penulis Korespondensi: [kbudiz@yahoo.com](mailto:kbudiz@yahoo.com)

**Abstrak**– Indonesia termasuk salah satu negara berkembang yang menghadapi kemiskinan khususnya di daerah kecil seperti Kelurahan Talang Babat yang terletak di Kabupaten Tanjung Jabung Timur Provinsi Jambi, penyebab terjadinya kemiskinan yaitu data warga penerima Bantuan Langsung Tunai tidak akurat dan tidak tepat sasaran. Hal ini dapat menimbulkan keluhan masyarakat, dan ketidakpuasan terhadap petugas kelurahan. Salah satu aspek penting yang mendukung strategi penanggulangan kemiskinan adalah dengan tersedianya data warga penerima Bantuan Langsung Tunai (BLT) yang akurat dan tepat sasaran. Oleh karena itu dibutuhkan metode untuk mengelompokkan data penerima BLT, agar pemberian BLT tepat sasaran. Pada penelitian ini algoritma *Naïve Bayes* diimplementasikan dan dianalisis menggunakan evaluasi 3 options dengan *Rapid miner* sebagai alat bantu penelitian. Dari hasil evaluasi menggunakan 3 options test yaitu *Use training set*, *5 Fold Cross-Validation* dan *10 Fold Cross-Validation*, didapatkan persentase nilai tertinggi pada pengujian *Use training set* yang akurasi sebesar 90,60%, presisi 96,67%, dan *recall* 92,06%.

**Kata Kunci:** *Data mining; Naïve Bayes; Gain ratio; Klasifikasi; Bantuan Langsung Tunai*

**Abstract**– Indonesia is one of the developing countries that faces poverty, especially in small areas such as Talang Babat Village which is located in East Tanjung Jabung Regency, Jambi Province, the cause of poverty is that data on residents who receive Direct Cash Assistance are inaccurate and not on target. This can lead to public complaints, and dissatisfaction with village officials. One important aspect that supports the poverty alleviation strategy is the availability of accurate and targeted data on recipients of Direct Cash Assistance (BLT). Therefore a method is needed to classify BLT beneficiary data, so that the provision of BLT is right on target. In this study the *Naïve Bayes* algorithm was implemented and analyzed using 3 options evaluation with *Rapid miner* as a research tool. From the results of the evaluation using the 3 options test, namely *Use training set*, *5 Fold Cross-Validation* and *10 Fold Cross-Validation*, the highest percentage score was obtained in the *Use training set* test with an accuracy of 90.60%, 96.67% precision, and 92% recall. .06%.

**Keywords:** *Data mining; Naïve Bayes; Gain ratio; Classification; Receiving Direct Cash Assistance*

### 1. PENDAHULUAN

Kemiskinan merupakan persoalan yang paling mendasar dan selalu menjadi pusat perhatian dunia. Menurut Laga Priseptian et al. [1] masalah utama bagi negara berkembang adalah kemiskinan, Indonesia juga termasuk negara berkembang yang menghadapi masalah kemiskinan. Banyak faktor yang mempengaruhi terjadinya kemiskinan, baik secara langsung atau tidak langsung. Menurut Itang [2] “Lapangan pekerjaan yang terbatas, kurangnya pendidikan, tingginya harga kebutuhan sehari-hari, akses sumber daya yang terbatas, pertumbuhan ekonomi yang tidak merata, produktifitas tenaga kerja, tingkat upah yang rendah, jenis pekerjaan, jumlah jam kerja, kesempatan kerja yang terbatas, dan inflasi, apalagi saat terjadi pandemi virus Covid-19 yang menyebabkan beberapa kebangkrutan serta terjadi pemutusan hubungan kerja yang mengakibatkan meningkatnya jumlah pengangguran dan kemiskinan.

Bantuan Langsung Tunai (BLT) diberikan oleh pemerintah kepada masyarakat miskin atau memiliki masalah ekonomi yang rendah. Pada Kelurahan Talang Babat Tanjung Jabung Timur banyak juga dijumpai masyarakat yang mempunyai masalah ekonomi dan tidak mendapatkan bantuan langsung tunai dari pemerintah daerah, padahal mereka layak mendapatkan bantuan tersebut, beberapa warga yang tidak mendapatkan bantuan ini menyuarkan masalahnya kepada pihak kelurahan agar mereka yang membutuhkan secara merata mendapatkan bantuan ini

Menurut Musfi Yendra et al. [3] “Bantuan Langsung Tunai (BLT) adalah program dari pemerintah dengan memberikan bantuan berupa uang tunai kepada masyarakat miskin agar kemiskinan di Indonesia berkurang, dengan adanya BLT ini diharapkan kemakmuran penduduk semakin merata”. Salah satu kesulitan yang dihadapi oleh pemerintah dalam pembagian bantuan langsung tunai ini adalah proses pembagian yang tidak merata dan tidak tepat sasaran, maka dibutuhkan pendataan yang lebih kompleks dan akurat terkait pihak mana saja yang berhak mendapatkan bantuan, untuk menentukannya diperlukan teknik *data mining*. Menurut Maulana [4], “*Data mining* merupakan proses penemuan pola data yang sudah ada didalam *database* dan dimanfaatkan untuk menyelesaikan masalah”.

Datamining telah banyak diterapkan untuk menyelesaikan masalah klasifikasi, salah satunya penerima bantuan sosial. Seperti pada penelitian yang dilakukan oleh Warisa mengenai klasifikasi penerima bantuan langsung tunai (BLT) dengan menerapkan metode *naïve bayes*. Penelitian ini menghasilkan, akurasi 83.67%,

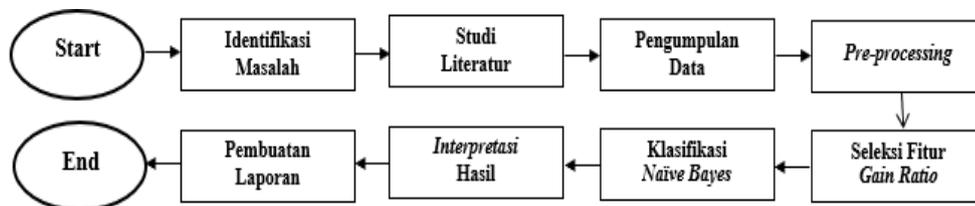
presisi 78.79% dan *recall* 74.29% [5]. Penelitian yang dilakukan oleh Amat Damuri, Umbar Riyanto, Hengki Rusdianto, dan Mohammad Aminudin, menerapkan metode *naïve bayes* dan *evaluasi confusion matrix* pada proses klasifikasi kelayakan penerima BLT di Bekasi. Dari hasil pengujian diperoleh nilai akurasi sebesar 86%, *recall* 85, dan presisi 88% [6]. Selanjutnya, penelitian Tohayah, Nur Ariesanto Ramdhan dan Herdian Bhakti menerapkan metode *Topsis* dan *Algoritma C4.5* dalam proses klasifikasi kelayakan penerima BLT. Pengujian yang dilakukan menghasilkan nilai akurasi sebesar 76,00%, *precision* sebesar 82,61%, dan *recall* sebesar 83,48% [7]. Sedangkan, penelitian Yosep Filki menerapkan Algoritma K-Means Clustering dalam proses klasifikasi data warga penerima BLT di Wali Nagari. Hasil pengujian memperoleh nilai akurasi sebesar 80,33% [8]. Pada penelitian Laila Qadrini, Andi Seppewali dan Asra Aina, metode *Decission Tree* dan *Adaboost* digunakan untuk klasifikasi tingkat penduduk miskin yang layak menerima Bantuan Langsung Tunai. Hasil pengujian menunjukkan nilai akurasi sebesar 80% dan 85% untuk menentukan penerima bantuan sosial secara tepat [9].

Berdasarkan hasil studi, metode *naïve bayes* menunjukkan performa yang sangat baik dalam mengklasifikasi data warga, selain itu merujuk nilai akurasi pada penelitian sebelumnya, terdapat peluang untuk meningkatkan performa klasifikasi data pada penelitian selanjutnya. Oleh karena itu, pada penelitian ini digunakan metode *naïve bayes* untuk klasifikasi penerima bantuan langsung tunai. Metode *naïve bayes* merupakan pengelompokan probabilistik sederhana yang melakukan perhitungan secara sistematis pada dataset yang diberikan [10]. Metode *naïve bayes* dapat mengasumsikan semua atribut independen yang diberikan oleh nilai pada variabel kelas [11]. Menurut Wahono [12], “Metode *naïve bayes* dinilai cukup sederhana dan mudah untuk diimplementasikan sehingga saat diuji dengan *dataset* yang benar maka metode ini akan sangat efektif”. Keuntungan menggunakan metode *naïve bayes* hanya membutuhkan sedikit *data training* untuk menentukan estimasi parameter dalam proses klasifikasi [13]. Untuk mengurangi *redundant* pada data, metode *naïve bayes* dapat dikombinasikan dengan pemilihan fungsi [14]. Metode *naïve bayes* memiliki akurasi yang tinggi dibandingkan metode lain dan memiliki performa yang bagus pada banyak domain [15]. Selain itu untuk meningkatkan akurasi klasifikasi pada pra-pemrosesan data dilakukan seleksi fitur. Seleksi fitur (*Feature Selection*) adalah salah satu teknik *data mining* yang umum digunakan pada tahapan *pre-processing* untuk memilih dan mengurangi fitur yang tidak relevan terhadap data, tujuannya untuk memilih fitur yang sesuai untuk pengujian [16]. Berdasarkan hasil studi, teknik seleksi fitur mampu meningkatkan performa klasifikasi dengan cara memilih fitur-fitur yang relevan. Teknik seleksi fitur yang digunakan pada penelitian ini adalah *gain ratio*. Menurut Sharma [17], “Metode *gain ratio* menggunakan teknik *scoring* untuk pembobotan sebuah fitur dengan menggunakan maksimal *entropy*, fitur yang dipilih memiliki nilai *gain ratio* yang lebih besar atau sama dengan nilai *threshold* tertentu.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Pada penelitian dibutuhkan kerangka kerja penelitian yang berisi langkah-langkah terstruktur agar penelitian dapat berjalan dengan baik. Pada Gambar 1 disajikan kerangka kerja penelitian.



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan Gambar 1, berikut adalah uraian dari langkah-langkah penelitian yang akan dilakukan :

a. Identifikasi Masalah

Pada tahap ini peneliti telah mengidentifikasi masalah yaitu beberapa warga Talang Babat yang kurang mampu tidak terdata atau tidak mendapatkan bantuan sosial berupa uang tunai dari pemerintah daerah.

b. Studi literatur

Pada tahap ini penulis mencari landasan teori sebagai referensi yang menguatkan penelitian, referensi didapatkan dari sumber yang valid seperti buku, jurnal, paper, dan lain sebagainya.

c. Pengumpulan data

Data yang digunakan dalam penelitian ini berjumlah 220 data warga Kelurahan Talang Babat Tajung Jabung Timur dan memiliki 12 atribut, diantaranya atribut Nama, L/P, Kelurahan, RT, Kepala Rumah Tangga, Status Penerima PKH, Status Penerima Rastra, Kondisi Rumah, Pekerjaan, Jumlah Penghasilan, Jumlah Tanggungan, Status

Penerima BLT. Pada tahap ini penulis menggunakan 2 metode pengumpulan data yaitu pengamatan langsung dan wawancara :

1. Pengamatan langsung (observation)

Penulis melakukan pengamatan secara langsung terhadap objek yang akan diteliti, tujuannya adalah untuk memperkuat data serta mengetahui informasi tentang kondisi warga yang layak mendapatkan bantuan langsung tunai.

2. Wawancara (interview)

Penulis melakukan wawancara dengan mengajukan pertanyaan secara langsung kepada pihak-pihak bersangkutan, misalnya Sekretaris Kelurahan Talang Babat Tanjung Jabung Timur serta staff atau karyawan yang mengelolah inputan data warga.

d. Pre-processing

pada tahap ini data mentah diproses dengan menggunakan *tools rapid miner* agar data yang *redundan* bisa dieliminasi

e. Seleksi Fitur *gain ratio*

pada tahap ini data yang telah melewati proses *cleaning* akan melalui tahap seleksi fitur yakni tahap menghapus fitur/atribut yang tidak relevan untuk proses perhitungan menggunakan metode *gain ratio*.

f. Klasifikasi *naïve bayes*

Data yang telah diseleksi fitur akan melalui tahap klasifikasi menggunakan metode *naïve bayes* dengan bantuan *tools rapid miner*.

g. Interpretasi Hasil

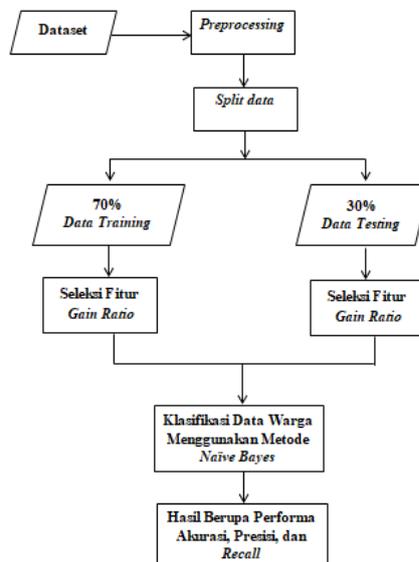
Pada tahap ini data dianalisis untuk mengevaluasi keakuratan data dengan metode *naïve bayes*, hasilnya disajikan sebagai nilai “Layak” dan “Tidak Layak”, dan validasi hasilnya diukur dengan nilai akurasi, presisi, *recall*.

h. Pembuatan laporan

Jika semua tahapan telah dilakukan maka akan disusun menjadi sebuah laporan penelitian yang lengkap, tujuannya adalah untuk dokumentasi dari penelitian yang telah dilakukan dan dapat dimanfaatkan pada waktu yang akan datang.

## 2.2 Alur Eksperimen

Untuk mempermudah proses penelitian, maka diperlukan alur eksperimen yang berisi tahapan-tahapan pengerjaan penelitian. Pada gambar 2 disajikan alur eksperimen pada penelitian.



Gambar 2. Alur Eksperimen

Berdasarkan Gambar 2, tahapan-tahapan pengerjaan penelitian yang diproses dengan menggunakan *tools rapid miner* adalah sebagai berikut :

a. Pre-processing

Data mentah diinput kedalam *rapid miner* yang akan diproses secara otomatis dan dapat mendeteksi data yang *redundant*, setelah itu dilakukan *cleaning* data untuk menghapus *noise* pada data.

b. Split Data

Data yang telah di *cleaning* akan di *split* dengan perbandingan 70% *data training* (Melatih algoritma) dan 30% *data testing* (Menguji algoritma)

c. Seleksi Fitur Gain Ratio

Setelah dilakukan *split data*, *data training* dan *data testing* akan melalui proses seleksi fitur dengan menggunakan metode *gain ratio* yang akan mengurangi atribut tidak *relevan* untuk perhitungan. *Gain ratio* membutuhkan perhitungan *Information Gain* terlebih dahulu karena *gain ratio* merupakan improvisasi dari perhitungan *information gain*. Sifat dari *gain ratio* adalah sebagai berikut [18] :

1. Bila data menyebar rata, *gain ratio* akan memiliki nilai besar.
2. Nilai akan menjadi kecil apabila semua data masuk dalam satu cabang.

Berikut persamaan gain ratio :

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)} \dots\dots\dots(1)$$

$$Gain(A) = - \sum_{i=1}^v (p_i * \log_{10} p_i) - \sum_{j=1}^v \left( \frac{|D_j|}{|D|} \right) * \left( - \sum_{j=1}^v (p_j * \log_{10} p_j) \right) \dots\dots\dots(2)$$

$$SplitInfo_A(D) = - \sum_{j=1}^v \left( \frac{|D_j|}{|D|} \right) * \log \left( \frac{|D_j|}{|D|} \right) \dots\dots\dots(3)$$

d. Klasifikasi Menggunakan Metode Naïve Bayes

Klasifikasi naïve bayes pada *tools rapid miner* menggunakan data training yang telah di seleksi fitur, lalu pengujian akan melalui 3 option test yaitu *use training set*, *5 fold cross validation*, dan *10 fold cross validation*. Berikut adalah persamaan dari naïve bayes [19] pada persamaan 4 :

$$P(C_i|X) = \frac{P(X|C_i)}{P(X)} \dots\dots\dots(4)$$

e. Hasil Evaluasi Metode Naïve Bayes

Setelah mengklasifikasi data warga Kelurahan Talang Babat Tanjung Jabung Timur menggunakan 3 option test, hasil performa dari klasifikasi diukur dengan tingkat akurasi, presisi, dan *recall*. Berikut adalah persamaan dari akurasi, presisi, dan *recall*.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \dots\dots\dots(5)$$

$$presisi = \frac{TP}{TP+FP} \dots\dots\dots(6)$$

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots(7)$$

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Pre-processing

Penelitian ini terdapat 220 data warga Kelurahan Talang Babat dan 12 atribut. Pada Tabel 1 disajikan profil data penelitian.

Tabel 1. Profil Data Penelitian

Kelas	Jumlah
Layak	176
Tidak Layak	44
Total	220

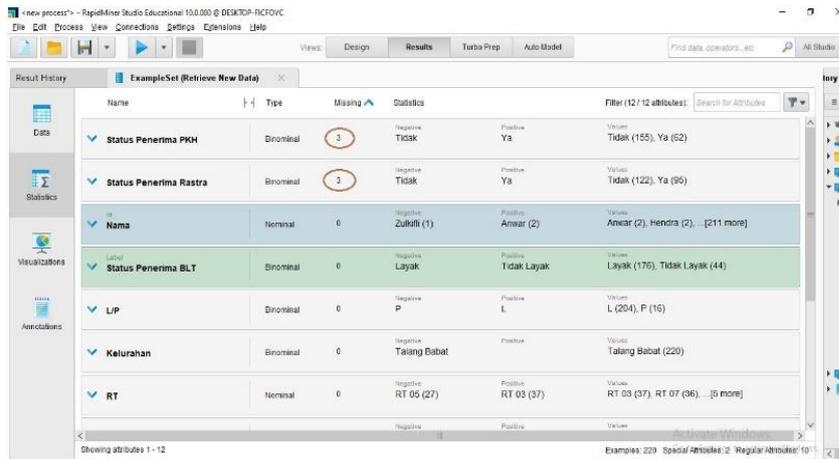
Berdasarkan hasil observasi yang telah dilakukan, penulis mendapatkan 12 atribut data yang akan dianalisis dan dilakukan perhitungan diantaranya terdapat atribut Nama, L/P, Kelurahan, RT, Kepala Rumah Tangga, Status Penerima PKH, Status Penerima Rastra, Kondisi Rumah, Pekerjaan, Jumlah Penghasilan, Jumlah Tanggungan, Status Penerima BLT. Pada Tabel 2 disajikan data warga Kelurahan Talang Babat penerima KPM.

Tabel 2. Data Warga Talang Babat Penerima KPM

No.	Nama	L/P	Kelurahan	RT	Kepala Rumah Tangga	Status Penerima PKH	Status Penerima Rastra	Kondisi Rumah	Pekerjaan	Jumlah Penghasilan	Jumlah Tanggungan	Status Penerima BLT
1	Sriyatin	P	Talang Babat	RT 01	Laki-Laki	Tidak	Tidak	Layak	Buruh	2600000	3	Layak
2	Supatmi	L	Talang Babat	RT 01	Perempuan	Tidak	Ya	Layak	ART	1500000	2	Layak
3	Badarun	L	Talang Babat	RT 01	Laki-Laki	Tidak	Ya	Layak	Karyawan Swasta	3500000	4	Layak
4	Rabuang	L	Talang Babat	RT 01	Laki-Laki	Tidak	Ya	Layak	Buruh	2600000	3	Layak

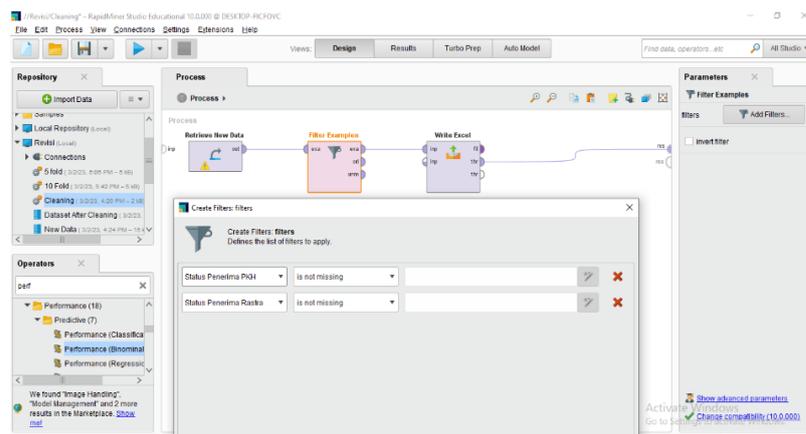
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
217	Faisal	L	Talang Babat	RT 07	Laki-Laki	Tidak	Tidak	Layak	Karyawan Swasta	3500000	3	Layak
218	Dodi Pomo	L	Talang Babat	RT 07	Laki-Laki	Ya	Tidak	Layak	Pedagang Kecil	1500000	2	Layak
219	Virhan	L	Talang Babat	RT 07	Laki-Laki	Tidak	Tidak	Layak	Buruh	2600000	4	Layak
220	Ripani	L	Talang Babat	RT 07	Laki-Laki	Tidak	Ya	Layak	Buruh	2600000	1	Tidak Layak

Data warga penerima KPM pada Tabel 2 diinput dengan cara memilih menu *import* data pada *rapid miner*, setelah mengatur datanya lalu atur area data yang digunakan jika sudah tekan tombol *next*, atur tipe data masing-masing atribut, lalu tambahkan *role id* pada atribut nama dan *role label* pada atribut status penerima BLT. Berdasarkan langkah-langkah yang telah dilakukan, Pada Gambar 3 disajikan statistik data yang mendeteksi adanya data kosong didalam *rapid miner*.



Gambar 3. Statistik Data

Pada gambar 3 ditemukan 6 data kosong, terdapat 3 dari atribut status penerima PKH dan 3 dari status penerima rastra. Untuk membersihkan/*cleaning* data kosong tersebut harus menggunakan dataset yang telah diinput kedalam *rapid miner*, operator filter example, serta write excel untuk menyimpan data yang telah diproses kedalam bentuk excel, Pada Gambar 4 disajikan alur data cleaning didalam *tools rapid miner*.

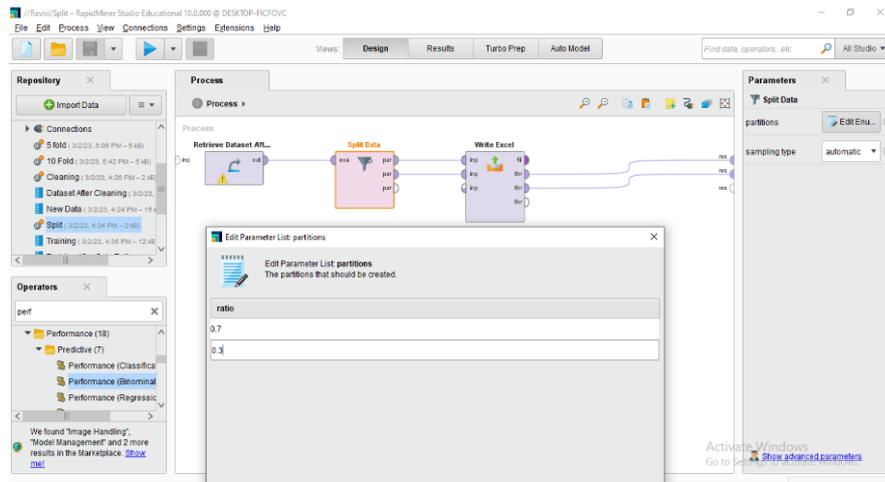


Gambar 4. Data Cleaning

Pada proses ini 6 data yang kosong akan dihapus otomatis dan menyisakan 214 data dari 220 data yang akan diproses pada tahap selanjutnya.

### 3.2 Split Data

Pada tahap ini data dibagi menjadi 2 dengan operator *split data*, 70% data training dan 30% data testing, proses disajikan pada Gambar 5.



Gambar 5. Alur Split Data

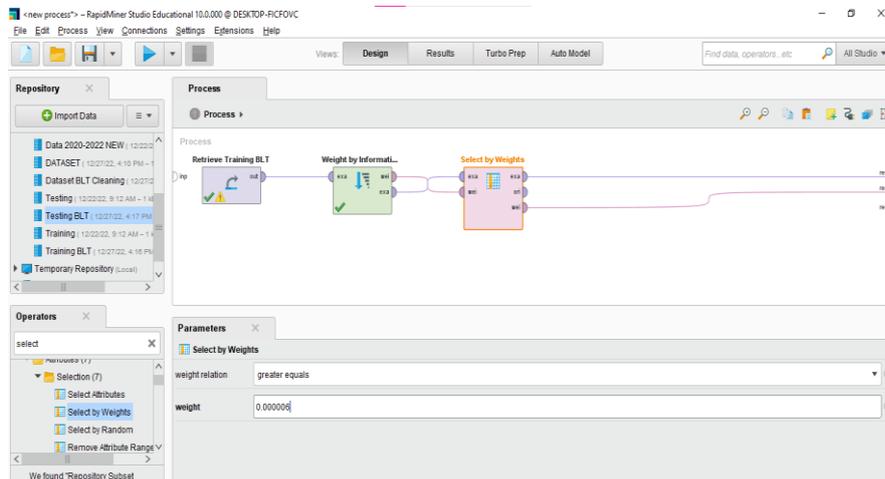
Setelah di *split data*, 70% data training menghasilkan 120 data “Layak” dan 29 data “Tidak Layak”, 30% data testing menghasilkan 52 data “Layak” dan 13 data “Tidak Layak”, yang disajikan pada Tabel 3.

Tabel 3. Hasil Split Data

Kelas	Training 70%	Testing 30%
Layak	120	52
Tidak Layak	29	13

### 3.3 Seleksi Fitur Gain Ratio

Pada tahap ini data yang tidak relevan untuk perhitungan dengan menggunakan data training, operator *weight by information gain* dan *select by weight* yang disajikan pada Gambar 6.



Gambar 6. Alur Seleksi Fitur Gain Ratio Data Training

Untuk menseleksi fitur dengan metode gain ratio, peneliti mengambil bobot  $>0,000002$  untuk batasan menseleksi atribut dan terdapat 7 atribut yang memiliki bobot  $>0,000002$ , Pada Tabel 4 disajikan 7 atribut hasil pembobotan dengan tools Rapid miner.

Tabel 4. Hasil Pembobotan Gain Ratio Data Training

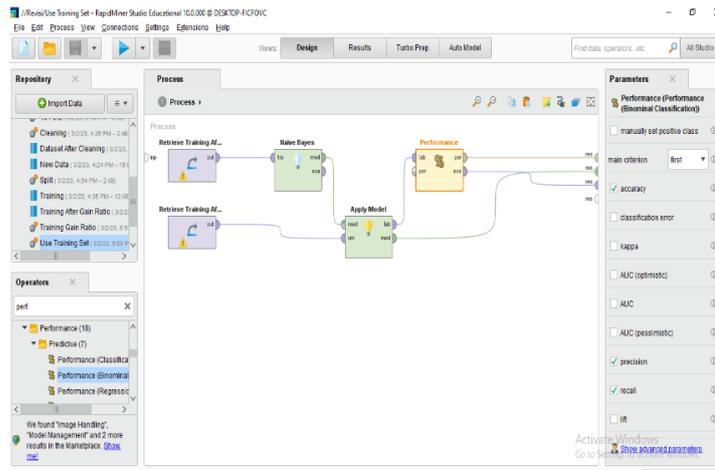
Atribut	Weight	Atribut	Weight
Kelurahan	0	RT	0.015
Kondisi Rumah	0	Kepala Rumah Tangga	0.025

Status Penerima Rastra	0.000002	Pekerjaan	0.053
Status Penerima PKH	0.002	Jumlah Penghasilan	0.078
L/P	0.010	Jumlah Tanggungan	0.255

3.4 Klasifikasi Menggunakan Metode *Naïve Bayes*

a. Pengujian Menggunakan *Use training set*

Untuk pengujian ini menggunakan *data training* yang telah diseleksi *fitur*, *naïve bayes*, *apply mode*, dan *performance* untuk melihat persentasi akurasi, presisi, *recall*. Pada Gambar 7 disajikan alur dari pengujian *use training set*.



Gambar 7. Alur *Use training set*

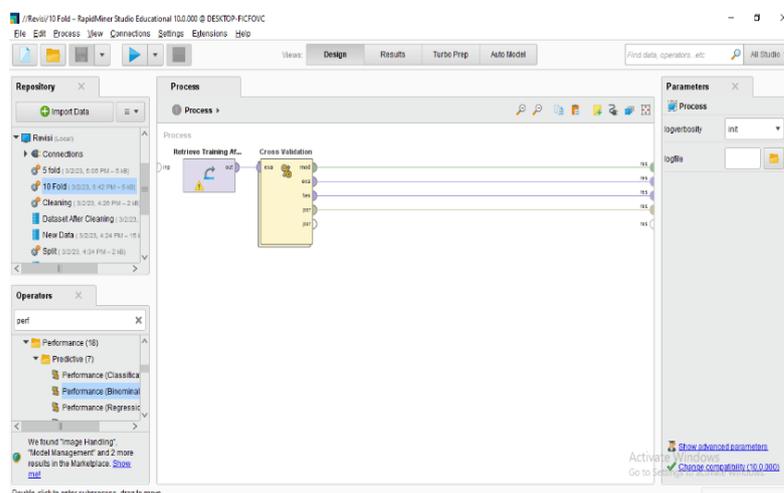
Berikut ini adalah hasil dari klasifikasi dari proses *Use training set* dengan menggunakan *data training* yang disajikan pada Tabel 5.

Tabel 5. Hasil Pengujian *Use training set*

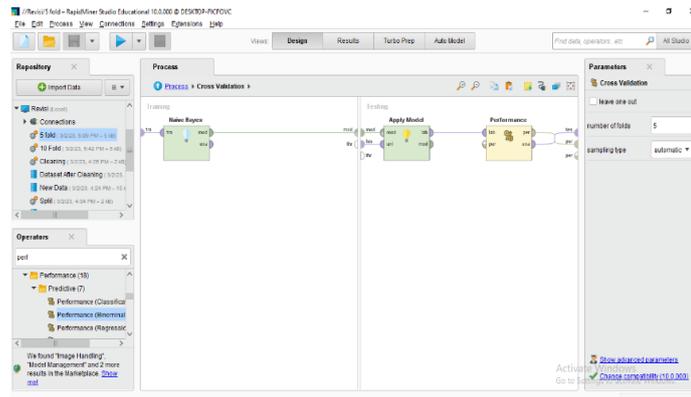
Accuracy	Recall ( <i>true.Layak</i> )	Recall ( <i>true.Tidak Layak</i> )	Precision ( <i>pred.Layak</i> )	Precision ( <i>pred.Tidak Layak</i> )
90,60%	96,67%	65,52%	92,06%	82.61%

b. Pengujian Menggunakan *5 fold cross validation*

Untuk pengujian ini menggunakan *data training* yang telah diseleksi *fitur*, *cross validation*, *naïve bayes*, *apply mode*, dan *performance* untuk melihat persentasi akurasi, presisi, *recall*, *number of fold* diisi dengan angka 5. Pada Gambar 8 dan 9 disajikan alur dari pengujian *5 fold cross validation*.



Gambar 8. Alur *5 Fold Cross Validation (1)*



Gambar 9. Alur 5 Fold Cross Validation (2)

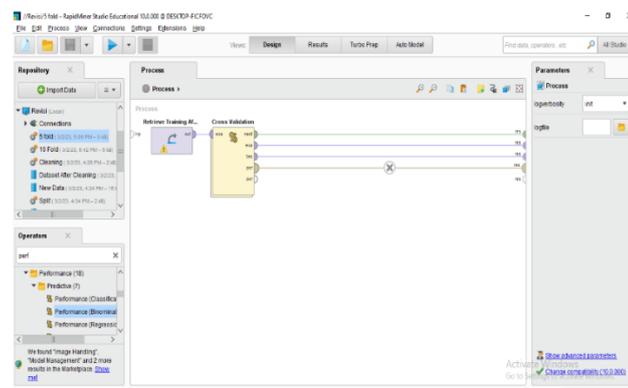
Berikut ini adalah hasil dari klasifikasi dari proses 5 fold cross validation (1) dan 5 fold cross validation (2) dengan menggunakan data training yang disajikan pada Tabel 6.

Tabel 6. Hasil Pengujian

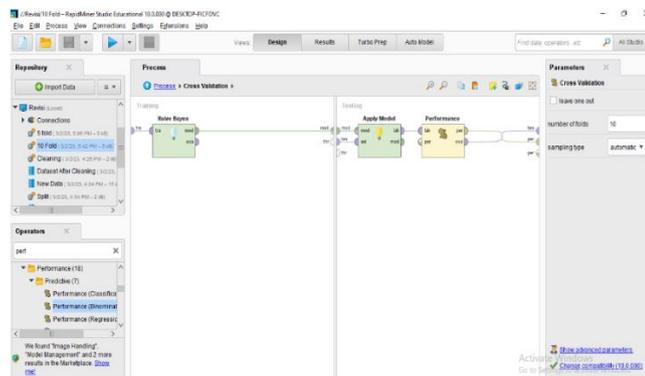
Accuracy	Recall (true.Layak)	Recall (true.Tidak Layak)	Precision (pred.Layak)	Precision (pred.Tidak Layak)
85,91%	95,00%	48,28%	88,37%	70.00%

c. Pengujian Menggunakan 10 fold cross validation

Pada pengujian ini cara perhitungannya sama dengan metode 5 fold cross validation yang membedakan hanya number of folds yang bernilai 10. Pada Gambar 10 dan 11 disajikan alur dari pengujian 5 fold cross validation.



Gambar 10. Alur 10 Fold Cross Validation (1)



Gambar 11. Alur 10 Fold Cross Validation (2)

Berikut ini adalah hasil dari klasifikasi dari proses 10 fold cross validation (1) dan 10 fold cross validation (2) dengan menggunakan data training yang disajikan pada Tabel 7.

Tabel 7. Hasil Pengujian

Accuracy	Recall (true.Layak)	Recall (true.Tidak Layak)	Precision (pred.Layak)	Precision (pred.Tidak Layak)
86,58%	94,17%	55,17%	89,68%	69,57%

### 3.5 Hasil Evaluasi Naïve Bayes

Setelah dilakukan analisis klasifikasi naïve bayes padg *tools rapid miner*. Berikut perbandingan hasil klasifikasi *naïve bayes* menggunakan 3 *option test* yang disajikan pada Tabel 8

Tabel 8. Model Evaluasi

Model Evaluasi	Akurasi	Presisi	Recall
Use training set	90,60%	96,67%	92,06%
5 Fold Cross Validation	85,91%	95,00%	88,37%
10 Fold Cross Validation	86,58%	94,17%	89,68%

Dari hasil evaluasi didapatkan akurasi tertinggi sebesar 90,60%, presisi 96,67%, dan *recall* 92,06% dengan pengujian *use training set*. berdasarkan pengujian *use training set* di *rapid miner* terdapat 13 keterangan penerima BLT pada data asli yang tidak sesuai dengan prediksi yang disajikan pada perhitungan menggunakan *tools rapid miner*, terdapat 9 data warga dengan status “Tidak Layak” mendapatkan bantuan pada data asli, menjadi “Layak” pada prediksi *rapid miner*, 4 data warga dengan status “Layak” mendapatkan bantuan pada data asli, menjadi “Tidak Layak” pada prediksi *rapid miner*.

## 4. KESIMPULAN

Penelitian ini menggunakan 220 data warga Kelurahan Talang Babat penerima KPM dan menggunakan 12 atribut diantaranya Nama, L/P, Kelurahan, RT, Kepala Rumah Tangga, Status Penerima PKH, Status Penerima Rastra, Kondisi Rumah, Pekerjaan, Jumlah Penghasilan, Jumlah Tanggungan, Status Penerima BLT. Dengan menggunakan metode gain ratio sebagai seleksi fitur bagi atribut yang tidak relevan, maka hasil akurasi dari penelitian ini memiliki kenaikan performa yang cukup signifikan dari penelitian sejenisnya. Proses pengujian akurasi, presisi, *recall* menggunakan 3 options yaitu *Use training set* dengan akurasi 90,60%, presisi 96,67%, dan *recall* 92,06%, 5 Fold Cross-Validation dengan akurasi 85,91% presisi 95,00%, *recall* 88,37%, dan 10 Fold Cross-Validation dengan akurasi 86,58%, presisi 94,17%, *recall* 89,68%. Adapun hasil test dengan menggunakan *Use training set* pada *Rapid miner* memiliki persentase nilai tertinggi dengan akurasi sebesar 90,60%, presisi 96,67%, dan *recall* 92,06%. Hasil dari class Layak berdasarkan confusional matrix sebanyak 126 data warga dari class Layak dan dari class Tidak Layak sebanyak 25 data. 5. Dari proses pengujian *use training set* menggunakan *tools rapid miner* didapatkan 13 hasil penerima bantuan langsung tunai (BLT) berbeda dengan prediksi *tools*, hal ini membuat *rapid miner* dapat memprediksi ketidak sesuaian dataa Berdasarkan hal tersebut dapat disimpulkan bahwa performa Naïve Bayes dengan seleksi fitur Gain Ratio dalam klasifikasi warga penerima bantuan langsung tunai memiliki hasil yang sangat baik.

## REFERENCES

- [1] L. Priseptian, W. Priana Primandhana, and F. Ekonomi dan Bisnis Universitas Pembangunan Nasional Veteran Jawa Timur, “Analisis faktor-faktor yang mempengaruhi kemiskinan,” *FORUM EKONOMI*, vol. 24, no. 1, pp. 45–53, 2022, [Online]. Available: <http://journal.feb.unmul.ac.id/index.php/FORUM EKONOMI>
- [2] F. Ekonomi and B. Islam, “FAKTOR FAKTOR PENYEBAB KEMISKINAN,” *Jurnal Keislaman, Kemasyarakatan dan Kebudayaan*, 2015.
- [3] N. Huda, M. Hasbi, and T. Susyanto, “Seleksi Penerima Bantuan Pangan Non Tunai di Desa Menggunakan Metode Naïve Bayes dan Simple Additive Weighting,” *Jurnal Ilmiah SINUS*, vol. 19, no. 1, p. 39, Jan. 2021, doi: 10.30646/sinus.v19i1.525.

- [4] C. Agus Sugianto and F. Rizky Maulana, "Algoritma Naïve Bayes Untuk Klasifikasi Penerima Bantuan Pangan Non Tunai ( Studi Kasus Kelurahan Utama ) Naive Bayes Algorithm For Classification of Non-Cash Food Aid Recipients (Case Study of Utama Village)," 2019.
- [5] A. Iphon Purnama, A. Aziz, A. Sartika Wiguna, and K. Kunci, "PENERAPAN DATA MINING UNTUK MENGLASIFIKASI PENERIMA BANTUAN PKH DESA WAE JARE MENGGUNAKAN METODE NAÏVE BAYES," *KURAWAL Jurnal Teknologi, Informasi dan Industri*, 2020, [Online]. Available: <https://jurnal.machung.ac.id/index.php/curawal>
- [6] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Riset Komputer)*, vol. 8, no. 6, p. 219, Dec. 2021, doi: 10.30865/jurikom.v8i6.3655.
- [7] Warisa and Nurahman, "Penerapan Metode Naive Bayes Dalam Mengklasifikasi Penerima BLT Pada Desa Pelangian," Bulan Oktober, 2022.
- [8] Anwar Pauji, S. Aisyah, A. Surip, R. Saputra, and I. Ali, "Implementasi Algoritma K-Nearest Neighbor Dalam Menentukan Penerima Bantuan Langsung Tunai," *KOPERTIP : Jurnal Ilmiah Manajemen Informatika dan Komputer*, vol. 4, no. 1, pp. 21–27, Jun. 2022, doi: 10.32485/kopertip.v4i1.114.
- [9] L. Qadrini, A. Seppewali, and A. Aina, "DECISION TREEDAN ADABOOSTPADA KLASIFIKASI PENERIMA PROGRAM BANTUAN SOSIAL," *Jurnal Inovasi Penelitian*, vol. 2, no. 7, 2021.
- [10] W. Supriyadi, Kusri, and A. Amborowati, "Perbandingan Kinerja Algoritma C4.5 Dan Naive Bayes Untuk Ketepatan Pemilihan Konsentrasi Mahasiswa," *Jurnal INFORMA Politeknik Indonusa Surakarta*, pp. 2442–7942, 2016.
- [11] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional," *Jurnal Tekno Kompak*, vol. 15, no. 1, pp. 131–145, 2021.
- [12] Wahono and Fitriyani, "Integrasi Bagging dan Greedy Forward Selection pada Prediksi Cacat Software dengan Menggunakan Naïve Bayes," *Jurnal Informatika*, vol. 1, no. 2, pp. 101–108, 2015.
- [13] S. A. Pattekari dan P. A., "Prediction System for Heart Disease Using Naive Bayes," *International Journal of Advanced Computer and Mathematical Sciences*, vol. 3, no. 3, pp. 2230–9624, 2012.
- [14] I. H. Witten, E. Frank, and M. A. Hall, "Data Mining Practical Machine Learning Tools and Techniques Elsevier," 2012.
- [15] Sugianto and T. H. Apandi, "Analisis Komparasi Machine Learning Pada Data Spam Sms," *J.TEDC*, vol. 12, no. 1, pp. 58–62, 2018.
- [16] I. Maulida, "Seleksi Fitur Dokumen Abstrak Menggunakan Metode Information Gain," *Jurnal Sifo Mikroskill*, vol. 7, no. 2, pp. 15–20, 2016.
- [17] A. Sharma and Dey, "Performance Investigation of Feature Selection Methods and Sentiment Lexicons for Sentiment Analysis," pp. 15–20, 2012.
- [18] R. H. Sukarna and Y. Ansori, "Implementasi Data Mining Menggunakan Metode Naive Bayes Dengan Feature Selection Untuk Prediksi Kelulusan Mahasiswa Tepat Waktu," *Jurnal Ilmiah Sains dan Teknologi*, vol. 6, no. 1, 2022.
- [19] D. Sartika and D. I. Sensuse, "Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian," vol. 1, no. 2, 2017.