

# Klasifikasi Kelayakan Penerima Bantuan Langsung Tunai Bagi UMKM di Kota Cimahi Menggunakan Algoritma *Naïve Bayes*

Castaka Agus Sugianto<sup>1</sup>, Putri Nurlaela Sari<sup>2</sup>

<sup>1,2,3,4</sup> Program Studi Teknik Informatika - Politeknik TEDC Bandung

Jl. Politeknik – Pesantren KM2 Cibabat Cimahi Utara – Cimahi Jawa Barat – Indonesia

[castaka@poltektedc.ac.id](mailto:castaka@poltektedc.ac.id), [putrinurlaela1328@gmail.com](mailto:putrinurlaela1328@gmail.com)

**Abstrak**— Dinas Sosial Kota Cimahi merupakan bagian dari instansi pemerintahan di Kota Cimahi. Dinas Sosial melaksanakan program yang ditetapkan pemerintah, yaitu program Bantuan Langsung Tunai untuk UMKM. Dalam pelaksanaan program ini, banyak pelaku UMKM yang mengeluh karena tidak menerima bantuan, sementara beberapa pelaku UMKM lainnya dianggap memiliki usaha besar dan pendapatan yang cukup justru mendapatkan bantuan. Oleh karena itu, penelitian ini melakukan pengolahan informasi dengan menggunakan *data mining* untuk mengelompokkan Kelayakan Penerima Bantuan Langsung Tunai bagi UMKM di Kota Cimahi dengan menggunakan Algoritma *Naïve Bayes*. Diharapkan hasil dari proses ini dapat digunakan sebagai bahan evaluasi bagi pemerintah. Pada penelitian ini, peneliti mengkategorikan data penerima bantuan langsung tunai UMKM menjadi “Layak” dan “Tidak Layak” dengan menggunakan teknik klasifikasi dalam *data mining* melalui Algoritma *Naïve Bayes*. Model *data mining* dihasilkan menggunakan *RapidMiner*, dengan nilai Probabilitas untuk kelas “Layak” sebesar 0,485 yang dibulatkan menjadi 0,48, sedangkan nilai Probabilitas untuk kelas “Tidak Layak” adalah 0,515 yang dibulatkan menjadi 0,51. Dari total 202 data untuk *class* “Layak” sebanyak 98 data, dan *class* “Tidak Layak” sebanyak 104 data. Dengan melakukan 3 (tiga) opsi pengujian yaitu 1 (satu) kali uji coba, 5 *fold cross validation*, dan 10 *fold cross validation*, algoritma *Naïve Bayes* menunjukkan tingkat Akurasi 100%, Presisi 100%, *Recall* 100%, dan AUC 1.00. Oleh karena itu, dapat disimpulkan bahwa hasil pengujian dengan algoritma *Naïve Bayes* adalah signifikan.

**Kata Kunci**— Kelayakan, *data mining*, klasifikasi, algoritma *naïve bayes*

**Abstract**— The Cimahi City Social Service is part of the government agency in Cimahi City. The Social Service implements a program set by the government, namely the Direct Cash Assistance program for MSMEs. In the implementation of this program, many MSME actors complained because they did not receive assistance, while several other MSME actors were considered to have large businesses and sufficient income actually received assistance. Therefore, this study conducts information processing using *data mining* to group the Eligibility of Direct Cash Assistance Recipients for MSMEs in Cimahi City using the *Naïve Bayes* Algorithm. It is hoped that the results of this process can be used as evaluation material for the government. In this study, the researcher categorized the data of MSME direct cash assistance recipients into "Feasible" and "Not Feasible" by using

*classification techniques in data mining through the Naïve Bayes Algorithm. The data mining model was generated using RapidMiner, with a Probability value for the "Feasible" class of 0.485 rounded to 0.48, while a Probability value for the "Unfeasible" class of 0.515 rounded to 0.51. Out of a total of 202 data, 98 data for the "Feasible" class, and 104 data for the "Not Worthy" class. By conducting 3 (three) test options, namely 1 (one) trial, 5 fold cross validation, and 10 fold cross validation, the Naïve Bayes algorithm shows a level of 100% Accuracy, 100% Precision, 100% Recall, and 1.00 AUC. Therefore, it can be concluded that the results of the test with the Naïve Bayes algorithm are significant.*

**Keywords**— Eligibility, *data mining*, classification, *naïve bayes algorithm*

## I. PENDAHULUAN

Peran usaha kecil, mikro dan menengah di Indonesia sangat penting, banyak masyarakat yang menjadikan usaha sebagai sumber penghasilan. Usaha Mikro Kecil dan Menengah (UMKM) merupakan entitas bisnis yang bersifat mandiri dan dijalankan oleh individu atau organisasi di seluruh sektor ekonomi [1]. Pasca pandemi covid-19 pada tahun 2020, para pelaku usaha mengalami keterpurukan yang berdampak pada usaha mereka. Semenjak saat itu, pemerintah membuat program bagi para pelaku usaha. Tujuannya untuk meningkatkan kembali produktivitas dan mengembangkan usaha mereka agar tetap terus berjalan.

Pemerintah Indonesia membuat program diantaranya adalah bantuan langsung tunai atau biasa disebut BLT. Bantuan Langsung Tunai (BLT) merupakan inisiatif pemerintah yang memberikan dukungan berbentuk uang tunai kepada masyarakat yang kurang mampu, supaya kemiskinan di Indonesia dapat berkurang. Dengan adanya Bantuan Langsung Tunai ini, diharapkan kesejahteraan masyarakat semakin merata [2].

Program Bantuan Langsung Tunai ditujukan untuk pelaku usaha kecil dan menengah dengan memberikan sejumlah uang tunai kepada pelaku usaha yang terkena dampak pandemi covid-19. Saat itu pemerintah hanya memberlakukan program ini pasca pandemi covid-19 saja, tetapi bantuan tersebut diperpanjang hingga sekarang. Pemerintah melihat bahwa bantuan ini seharusnya tetap dilaksanakan untuk

meningkatkan dan mengembangkan usaha masyarakat. Dalam pelaksanaan program BLT-UMKM harus langsung menyentuh dan memberikan manfaat kepada pelaku usaha.

Program BLT-UMKM di Indonesia khususnya di Kota Cimahi, Jawa Barat tidak dapat maksimal dalam penyalurannya. Sasaran program BLT-UMKM di Kota Cimahi yang kurang tepat disebabkan karena ketidaksesuaian kriteria dalam penentuan kelayakan penerima bantuan langsung tunai bagi pelaku usaha.

Data yang akurat sangat diperlukan, karena ini akan menjadi salah satu cara untuk memudahkan pemerintah dalam menentukan kelayakan penerima bantuan. Maka dari itu perlu adanya sebuah metode yang dapat mengolah dan mengelompokkan dalam penentuan kelayakan penerima bantuan langsung tunai bagi UMKM dengan cepat dan sesuai. Salah satu teknik yang digunakan adalah teknik *Data Mining*.

*Data mining* adalah suatu teknik ekstraksi atau penggalian data yang belum diketahui sebelumnya, tetapi dapat dipahami dan bermanfaat dari *database* yang besar dan digunakan untuk membuat keputusan penting bagi bisnis [3]. Ada beberapa teknik dalam *data mining* seperti *classification*, *clustering*, *association*, *regression*, *forecasting*, *sequence analysis* dan *deviation analysis* [4].

Klasifikasi adalah metode untuk menemukan pola untuk membedakan kelas data yang berbeda satu sama lain dan untuk mengidentifikasi objek yang termasuk dalam kategori tertentu dengan melihat bagaimana kelompok tersebut bertindak dan karakteristiknya [5]. Salah satu metode yang sering digunakan adalah metode *Naïve Bayes Classifier*.

Teori probabilitas kondisional, atau algoritma Bayes dengan tingkat independensi yang tinggi pada data, adalah dasar dari metode prediksi probabilitas *Naïve Bayes Classifier*, yang berarti bahwa fitur tertentu dalam kumpulan data tertentu tidak terkait dengan fitur lain dalam kumpulan data yang sama [6].

Banyak penelitian terdahulu yang menggunakan metode *Naïve Bayes Classifier*, diantaranya penelitian yang dilakukan oleh Supriyanto [7] terkait Penerapan Algoritma *Naïve Bayes Classifier* untuk penentuan Bantuan Bedah Rumah di Kelurahan Krpyak hasilnya mendapatkan *Accuracy* sebesar 100%, *Precision* 100% dan *Recall* 100%. Penelitian lainnya yaitu penerapan Algoritma *Naïve Bayes* dalam Penentuan Kelayakan Penerima Bantuan Program Keluarga Harapan hasilnya mendapatkan akurasi sebesar 100% [8]. Sedangkan, yang sama dengan penelitian ini yaitu penelitian yang dilakukan oleh Sugianto [9] Algoritma *Naïve Bayes* untuk Klasifikasi Penerima Bantuan Pangan Non Tunai Di Kelurahan Utama mendapatkan *Accuracy* sebesar 58,29%, *Precision* 92,90%, *Recall* 21,84% dan AUC 0,765.

Hasil dari penelitian ini diharapkan dapat menjadi informasi bagi pemerintah Kota Cimahi ataupun pihak yang mengurus dibidang UMKM untuk dijadikan bahan evaluasi terhadap penentuan kelayakan penerima bantuan langsung tunai bagi pelaku UMKM lebih baik dan sesuai dengan apa yang menjadi targetan pemerintah dalam penyaluran program bantuan dana.

## II. METODE PENELITIAN

Penelitian ini mengajukan konsep model yang menggunakan algoritma *Naïve Bayes* untuk mengklasifikasi kelayakan penerima bantuan langsung tunai bagi UMKM di Kota Cimahi. Metodologi penelitian melibatkan serangkaian proses sebagai berikut:

### A. Menentukan Masalah

Menentukan masalah adalah proses awal peneliti. Proses ini sudah dijelaskan sebelumnya tentang klasifikasi kelayakan penerima bantuan langsung tunai bagi UMKM di kota Cimahi menggunakan algoritma *Naïve Bayes*.

### B. Menentukan Tujuan dan Ruang Lingkup

Menentukan tujuan dan ruang lingkup merupakan apa yang akan dicapai dalam penelitian ini, serta batasan yang dilakukan agar penelitian tidak terlalu luas dan menjadi lebih efektif.

### C. Mencari Literatur

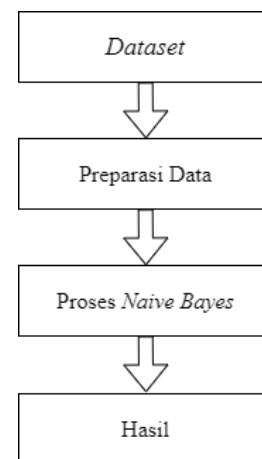
Dilakukan pencarian literatur atau referensi untuk menemukan berbagai sumber terkait data mining, metode klasifikasi, dan algoritma *naive bayes*.

### D. Pengumpulan Data

Pengumpulan data dilakukan dengan cara observasi ke bagian dinas sosial Kota Cimahi dan mengumpulkan data dari hasil pendataan lengkap KUMKM pada tahun 2023.

### E. Metode yang Diusulkan

Metode yang diusulkan dijelaskan proses pengolahan data pada penelitian, tahapan tersebut dimulai dari *dataset*, persiapan data, proses dan hasil dari algoritma *naive bayes*.



Gbr. 1 Metode yang Diusulkan

#### 1) Dataset Pelaku UMKM Kota Cimahi

*Dataset* yang digunakan dalam penelitian ini adalah data pelaku UMKM yang diperoleh langsung dari Dinas Sosial Kota Cimahi setelah mendapatkan persetujuan dari pihak yang bersangkutan. Data tersebut mencakup total 273 *record* yang berasal dari pendataan PL-KUMKM tahun 2023. Data yang diperoleh terdiri dari 8 *field* yaitu nama pelaku usaha, jenis

kelamin, pendidikan, pekerjaan, jumlah usaha, jenis usaha, pendapatan usaha dan keterangan. Pada Tabel 1 terdapat *dataset* yang digunakan.

TABEL I  
SAMPEL DATA PELAKU UMKM

No	Nama Pelaku Usaha	Jenis Kelamin	Pendidikan	Pekerjaan	Jumlah Usaha	Jenis Usaha	Pendapatan Usaha	Keterangan
1.	P1	Laki-laki	SLTA	Tidak	1	Kecil	1500000	Layak
2.	P2	Laki-laki	SD	Ya	1	Kecil	1500000	Layak
3.	P3	Laki-laki	SLTA	Tidak	1	Menengah	4500000	Tidak Layak
4.	P4	Laki-laki	SLTA	Tidak	1	Menengah	2500000	Tidak Layak
5.	P5	Perempuan	S1	Ya	2	Menengah	3500000	Tidak Layak
6.	P6	Perempuan	SLTA	Ya	1	Menengah	2000000	Tidak Layak
7.	P7	Laki-laki	S1	Ya	1	Menengah	5000000	Tidak Layak
8.	P8	Perempuan	SLTA	Tidak	1	Kecil	1500000	Layak
9.	P9	Laki-laki	SMP	Tidak	1	Kecil	1500000	Layak
10.	P10	Laki-laki	SLTA	Tidak	1	Menengah	2000000	Tidak Layak
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
273.	P273	Laki-laki	D3	Ya	1	Menengah	2200000	Tidak Layak

2) *Preparasi Data*

Setelah mendapatkan data, langkah selanjutnya adalah melakukan preparasi atau pembersihan data. Ini dilakukan untuk menghilangkan informasi yang tidak penting untuk pengolahan lebih lanjut. Tujuannya adalah agar data menjadi bersih dan siap untuk dianalisis lebih lanjut. Jumlah *dataset* dalam penelitian ini terdiri dari 273 dengan 8 field. Namun, sebanyak 71 *record* dihapus karena mengandung informasi yang tidak lengkap sehingga tidak diperlukan untuk proses *data mining*. Beberapa data yang dihapus mencakup informasi tentang nama pelaku usaha, jenis kelamin dan pendidikan. Penulis melakukan beberapa modifikasi untuk mempermudah proses *data mining*. Sehingga dalam penelitian ini data yang digunakan sebanyak 202 dengan 5 *field* yang terdiri dari pekerjaan, jumlah usaha, jenis usaha, pendapatan usaha dan keterangan. Pada Tabel 2 terdapat preparasi data yang digunakan.

TABEL II  
PREPARASI DATA

No	Pekerjaan	Jumlah Usaha	Jenis Usaha	Pendapatan Usaha	Keterangan
1.	Tidak	1	Kecil	1500000	Layak
2.	Ya	1	Kecil	1500000	Layak
3.	Tidak	1	Menengah	4500000	Tidak Layak
4.	Tidak	1	Menengah	2500000	Tidak Layak
5.	Ya	2	Menengah	3500000	Tidak Layak
6.	Ya	1	Menengah	2000000	Tidak Layak
7.	Ya	1	Menengah	5000000	Tidak Layak
8.	Tidak	1	Kecil	1500000	Layak
9.	Tidak	1	Kecil	1500000	Layak
10.	Tidak	1	Menengah	2000000	Tidak Layak
.	.	.	.....	.	.
.	.	.	.....	.	.
.	.	.	.....	.	.
202.	Ya	1	Menengah	2200000	Tidak Layak

F. *Proses dan Hasil Penelitian*

Dengan menggunakan algoritma *Naive Bayes*, penulis akan melakukan klasifikasi data penerima BLT-UMKM dan menggunakan hasil dari proses *RapidMiner* untuk menghasilkan nilai kemungkinan dan akurasi dari variabel yang ditargetkan dalam proses pengalihan data.

G. *Analisis*

Analisis ini dilakukan dengan mengamati dan menginterpretasikan hasil yang diperoleh *RapidMiner*. Hasil analisis ini akan menghasilkan nilai kemungkinan dan akurasi dari data yang diterima.

H. *Penarikan Kesimpulan*

Penarikan kesimpulan disini sebagai salah satu bagian dari proses penelitian, di mana penulis mencari makna umum dari informasi yang diperoleh.

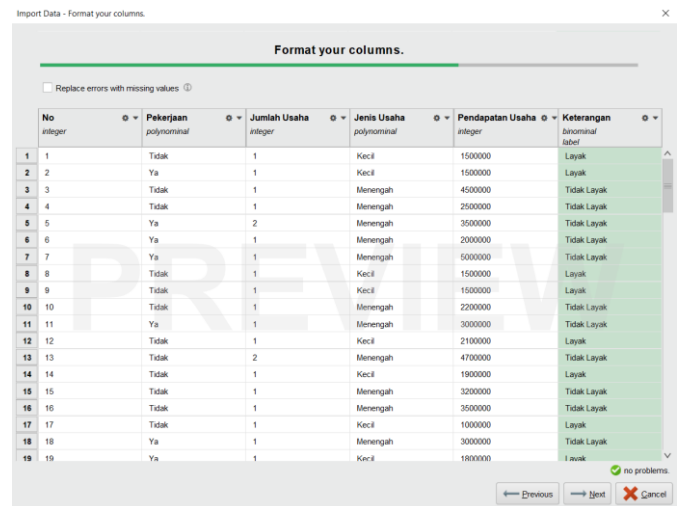
III. HASIL DAN PEMBAHASAN

Pada bagian ini merupakan proses dimulainya pengujian pada data dengan menggunakan *RapidMiner*. Berikut langkah-langkah melakukan pengujian *naive bayes classifier* pada *RapidMiner*:

A. *Proses Pengujian Naive Bayes Classifier pada RapidMiner*

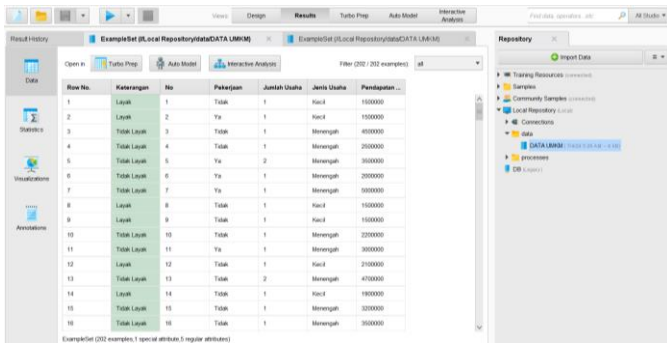
1) *Import Data*

Pada langkah ini dilakukan *import* data yang akan diuji dan mengubah *type* menjadi *binomial* apabila kolom hanya berisi 2 (dua) karakter, *integer* untuk kolom yang berisi nomor saja mengubah *role* pada kolom kondisi menjadi *label*. Atribut ini berperan sebagai atribut target untuk operator pembelajaran atau sering disebut dengan “variabel target”. Proses *import* data dapat dilihat pada Gambar 1.



Gbr. 2 Proses *Import* Data

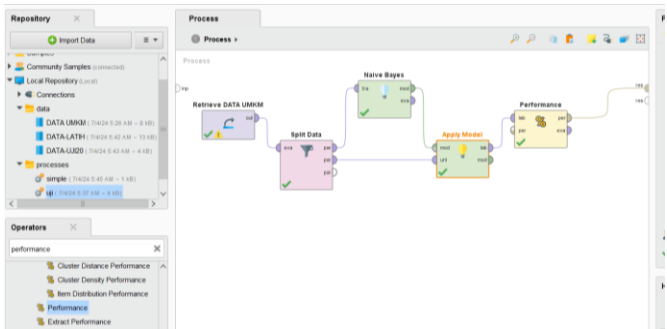
Pada Gambar 3 disajikan hasil *import* data yang telah mendeteksi sebanyak 202 data dan selanjutnya akan dilakukan proses *split* data.



Gbr. 3 Hasil Import Data

2) Split Data

Dalam proses ini, data ditambahkan ke lembar kerja dengan *drag & drop* atau *copy paste*, dan kemudian dibagi menjadi dua dengan menggunakan operator *split data*. 20% untuk data pengujian dan 80% untuk data pelatihan. Ditambahkan operator *Naive Bayes* untuk dihubungkan dengan data dan operator *Apply Model* yang berfungsi untuk mengecek data dari model *Naive Bayes*, setelah itu ditambahkan operator *performance* yang berfungsi untuk mengetahui sebaran data serta akurasi dari data yang telah diujikan. Proses *split data* dapat dilihat pada Gambar 4.



Gbr. 4 Proses Split Data

Pada pengolahan data kelayakan penerima bantuan langsung tunai bagi UMKM Kota Cimahi menggunakan algoritma *Naive Bayes* dilakukan pengujian sebanyak 1 (satu) kali, penulis mendapatkan hasil dari 5 atribut dengan nilai pred.Layak sebanyak 20 data dan nilai pred.Tidak Layak sebanyak 21 data dengan akurasi pengujian sebesar 100%. Hasil akurasi dapat dilihat pada Gambar 5.

Criterion	accuracy	precision	recall	AUC (optimistic)	AUC (pessimistic)
accuracy	100.00%				
precision					
recall					
AUC (optimistic)					
AUC (pessimistic)					

Gbr. 5 Hasil Akurasi Algoritma Naive Bayes

Selain hasil akurasi yang didapat, adapun hasil dari Presisi pada pengujian ini sebesar 100%. Presisi adalah proporsi prediktif positif artinya dari semua prediksi yang dikategorikan sebagai “Layak” dan “Tidak Layak” dikatakan benar. Hasil presisi dapat dilihat pada Gambar 6.

Criterion	accuracy	precision	recall	AUC (optimistic)	AUC (pessimistic)
accuracy	100.00%				
precision					
recall					
AUC (optimistic)					
AUC (pessimistic)					

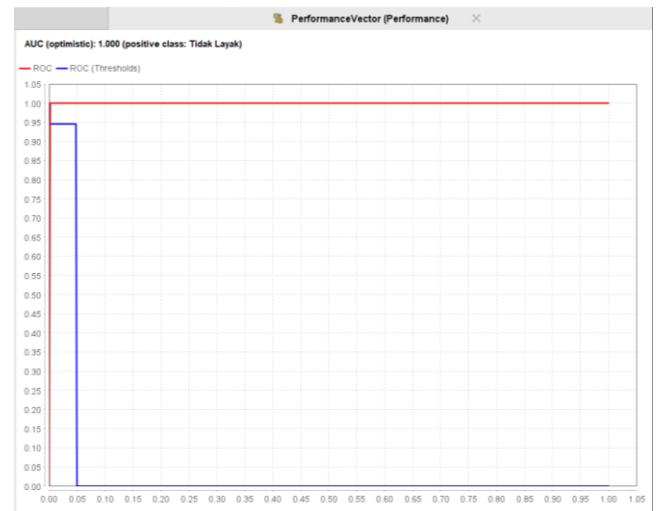
Gbr. 6 Hasil Presisi Algoritma Naive Bayes

Adapun hasil dari *Recall* pada pengujian ini sebesar 100%. Metode evaluasi yang digunakan untuk mengukur kemampuan model untuk menemukan semua contoh yang relevan dalam data dikenal sebagai *recall*. Hasil *recall* dapat dilihat pada Gambar 7.

Criterion	accuracy	precision	recall	AUC (optimistic)	AUC (pessimistic)
accuracy	100.00%				
precision					
recall					
AUC (optimistic)					
AUC (pessimistic)					

Gbr. 7 Hasil Recall Algoritma Naive Bayes

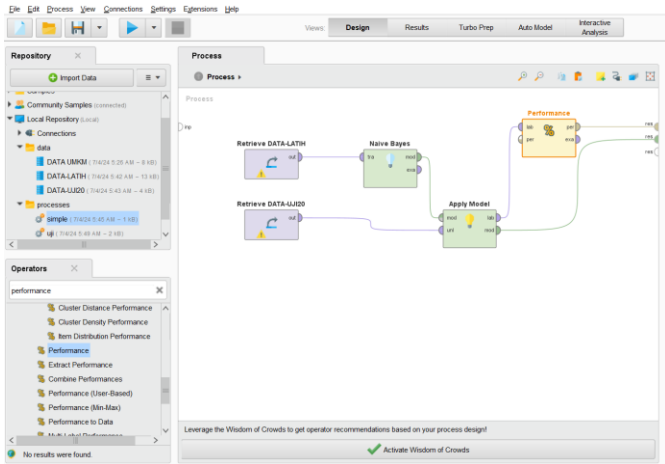
Pada Gambar 8 menunjukkan grafik AUC (*Area Under Curve*) dan kurva ROC (*Receiver Operating Characteristic*) untuk sebuah model klasifikasi. AUC yang dihasilkan sebesar 1.000 yang dapat diartikan dengan 1.00, ini menunjukkan bahwa model memiliki kemampuan prediksi yang sempurna tanpa ada kesalahan.



Gbr. 8 Grafik AUC (Area Under Curve) Algoritma Naive Bayes

B. Simple Distribution Naive Bayes

Pada tahap ini dilakukan penambahan operator *Naive Bayes* untuk dihubungkan dengan data latih, setelah itu penulis menambahkan operator *Apply Model* yang berfungsi untuk mengecek data latih dari model *Naive Bayes* dan dihubungkan dengan data *testing* yang dimana data *testing* ini berfungsi sebagai perbandingan data latih yang ada pada *Apply Model*, setelah itu ditambahkan operator *performance* yang berfungsi untuk mengetahui sebaran data serta akurasi dari data latih yang telah diujikan menggunakan data *testing*. Proses pengujian tersebut dapat dilihat pada Gambar 9.



Gbr. 9 Proses Pengujian Menggunakan Algoritma Naive Bayes

Nilai probabilitas dihitung dengan menggunakan hasil model distribusi algoritma *naive bayes* dari proses data mining. Nilai untuk atribut keterangan nilai *class* "Layak" sebesar 0,485 dan nilai untuk atribut keterangan nilai *class* "Tidak Layak" sebesar 0,515 ditunjukkan pada Gambar 10.

### Simple Distribution

Distribution model for label attribute Keterangan

Class Layak (0.485)  
5 distributions

Class Tidak Layak (0.515)  
5 distributions

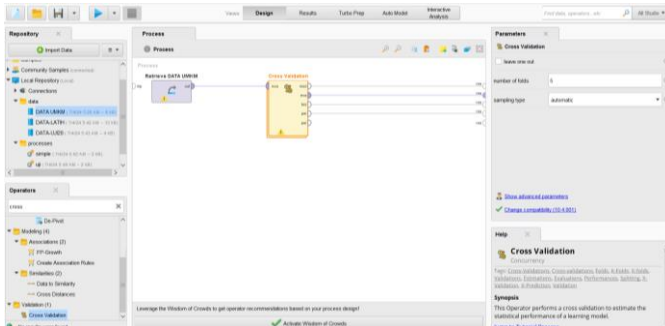
Gbr. 10 Hasil Simple Distribution Algoritma Naive Bayes

### C. Proses Validasi menggunakan Cross Validation

*Cross Validation* adalah teknik yang digunakan untuk mengevaluasi model *machine learning* agar dapat mengukur performa model secara akurat. Pengujian dilakukan dengan menggunakan jumlah *fold* yang berbeda yaitu:

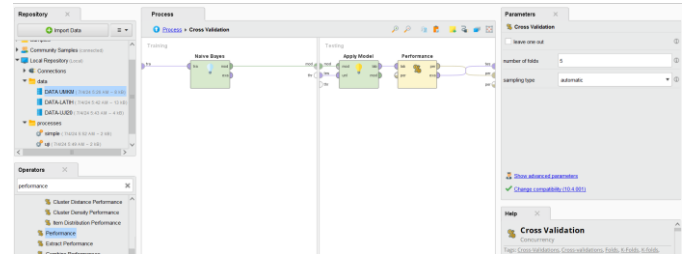
#### 1) Pengujian menggunakan 5 Fold Cross Validation

Teknik ini akan dilakukan pengujian sebanyak 5 kali agar menghasilkan nilai yang lebih optimal. Berikut adalah hasil pengujian *5 fold cross validation* yang dapat dilihat pada Gambar 11.



Gbr. 11 Proses Cross Validation 5 fold

Pada pengujian ini menggunakan data fitur, *cross validation*, *naive bayes*, penerapan model, dan kinerja untuk mengukur persentase akurasi, presisi, *recall*, dan jumlah *fold* yang diisi dengan angka 5. Proses *5 fold cross validation* dapat dilihat pada Gambar 12.



Gbr. 12 Proses Algoritma Di Dalam Cross Validation 5 fold

Berikut adalah hasil pengujian menggunakan Algoritma *Naive Bayes* dan *Cross Validation*. Tab *PerformanceVector* menampilkan berbagai metrik untuk mengevaluasi kinerja model klasifikasi. Hasil yang didapat dari 5 atribut dengan nilai pred.Layak sebanyak 98 dan nilai pred.Tidak Layak sebanyak 104 dengan akurasi pengujian sebesar 100%. Hasil akurasi dapat dilihat pada Gambar 13.

Criterion	Table View	Plot View
accuracy	accuracy: 100.00% +/- 0.00% (micro average: 100.00%)	
precision		
recall		
AUC (optimistic)		
AUC		
AUC (pessimistic)		

	true Layak	true Tidak Layak	class precision
pred. Layak	98	0	100.00%
pred. Tidak Layak	0	104	100.00%
class recall	100.00%	100.00%	

Gbr. 13 Hasil Akurasi 5 fold Cross Validation

Adapun hasil dari Presisi pada pengujian ini sebesar 100%. Hasil presisi dapat dilihat pada Gambar 14.

Criterion	Table View	Plot View
accuracy	accuracy: 100.00% +/- 0.00% (micro average: 100.00%)	
precision	precision: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: Tidak Layak)	
recall		
AUC (optimistic)		
AUC		
AUC (pessimistic)		

	true Layak	true Tidak Layak	class precision
pred. Layak	98	0	100.00%
pred. Tidak Layak	0	104	100.00%
class recall	100.00%	100.00%	

Gbr. 14 Hasil Presisi 5 fold Cross Validation

Hasil dari *Recall* pada pengujian ini sebesar 100%. Hasil *recall* dapat dilihat pada Gambar 15.

Criterion	Table View	Plot View
accuracy	accuracy: 100.00% +/- 0.00% (micro average: 100.00%)	
precision	precision: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: Tidak Layak)	
recall	recall: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: Tidak Layak)	
AUC (optimistic)		
AUC		
AUC (pessimistic)		

	true Layak	true Tidak Layak	class precision
pred. Layak	98	0	100.00%
pred. Tidak Layak	0	104	100.00%
class recall	100.00%	100.00%	

Gbr. 15 Hasil Recall 5 fold Cross Validation

Hasil AUC-1.00 dari analisis data di atas menggunakan RapidMiner dengan pengukuran Naive Bayes menunjukkan bahwa data berada dalam kategori sangat baik (klasifikasi baik) [10]. Tingkat akurasi AUC dibagi menjadi lima kategori, yaitu:

- 0.90 – 1.00 = Klasifikasi Sangat Baik
- 0.80 – 0.90 = Klasifikasi Baik
- 0.70 – 0.80 = Klasifikasi Rata-rata
- 0.60 – 0.70 = Klasifikasi Buruk

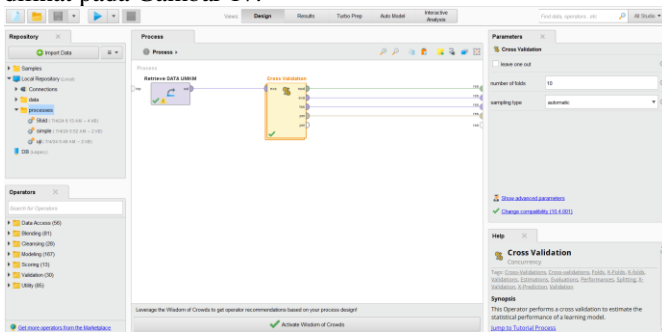
- $0.50 - 0.60 = \text{Kegagalan}$



Gbr. 16 Grafik AUC (Area Under Curve) 5 fold Cross Validation

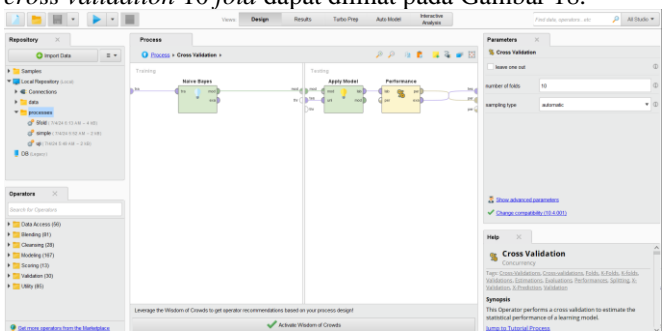
2) Pengujian menggunakan 10 Fold Cross Validation

Cara perhitungan untuk pengujian ini sama dengan yang dilakukan sebelumnya, yaitu 5 fold cross validation. Yang membedakan hanya jumlah fold yang bernilai 10, yang dapat dilihat pada Gambar 17.



Gbr. 17 Proses Cross Validation 10 fold

Pengujian ini menggunakan data yang dipilih untuk fitur, cross validation, naive bayes, penerapan model, dan kinerja. Tujuannya adalah untuk mengukur persentase akurasi, presisi, recall, dan number of fold yang diisi dengan angka 10. Proses cross validation 10 fold dapat dilihat pada Gambar 18.



Gbr. 18 Proses Algoritma Di Dalam Cross Validation 10 fold

Hasil pengujian yang dilakukan menggunakan Algoritma Naive Bayes dan Cross Validation ditunjukkan di sini. Penulis mendapatkan hasil dari 5 atribut dengan nilai pred.Layak sebanyak 98 dan nilai pred.Tidak Layak sebanyak 104 dengan

akurasi pengujian sebesar 100%. Hasil akurasi dapat dilihat pada Gambar 19.

Criterion	accuracy	precision	recall	AUC (optimistic)	AUC	AUC (pessimistic)
accuracy	100.00% +/- 0.00% (micro average: 100.00%)					
AUC (optimistic)						
AUC						
AUC (pessimistic)						
pred. Layak	98	0	100.00%			
pred. Tidak Layak	0	104	100.00%			
class recall	100.00%	100.00%				

Gbr. 19 Hasil Akurasi 10 fold Cross Validation

Adapun hasil dari Presisi pada pengujian ini sebesar 100%. Hasil presisi dapat dilihat pada Gambar 20.

Criterion	precision	recall	AUC (optimistic)	AUC	AUC (pessimistic)
precision	100.00% +/- 0.00% (micro average: 100.00%) (positive class: Tidak Layak)				
AUC (optimistic)					
AUC					
AUC (pessimistic)					
pred. Layak	98	0	100.00%		
pred. Tidak Layak	0	104	100.00%		
class recall	100.00%	100.00%			

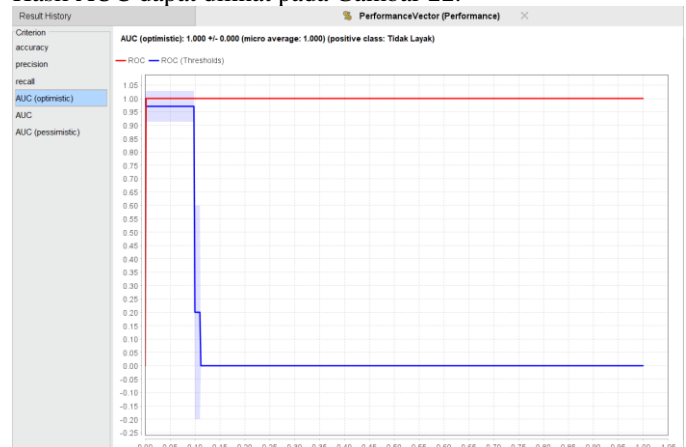
Gbr. 20 Hasil Presisi 10 fold Cross Validation

Hasil dari Recall pada pengujian ini sebesar 100%. Hasil recall dapat dilihat pada Gambar 21.

Criterion	recall	precision	AUC (optimistic)	AUC	AUC (pessimistic)
recall	100.00% +/- 0.00% (micro average: 100.00%) (positive class: Tidak Layak)				
AUC (optimistic)					
AUC					
AUC (pessimistic)					
pred. Layak	98	0	100.00%		
pred. Tidak Layak	0	104	100.00%		
class recall	100.00%	100.00%			

Gbr. 21 Hasil Recall 10 fold Cross Validation

Dari analisis data yang disebutkan di atas dengan RapidMiner dan pengukuran Naive Bayes, hasil AUC 1.00. Hasil AUC dapat dilihat pada Gambar 22.



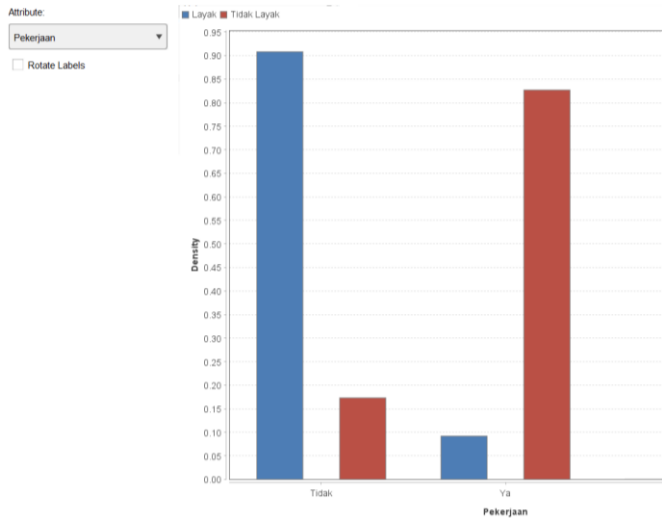
Gbr. 22 Grafik AUC (Area Under Curve) 10 fold Cross Validation

D. Kelayakan Penerima BLT-UMKM Berdasarkan Atribut

Pada Gambar 23 menunjukkan sebuah histogram distribusi kelayakan penerima bantuan langsung tunai bagi UMKM berdasarkan atribut "Pekerjaan".

Dengan demikian, pekerjaan memiliki pengaruh yang signifikan terhadap kelayakan penerima BLT-UMKM. Dimana, mereka yang tidak memiliki pekerjaan cenderung

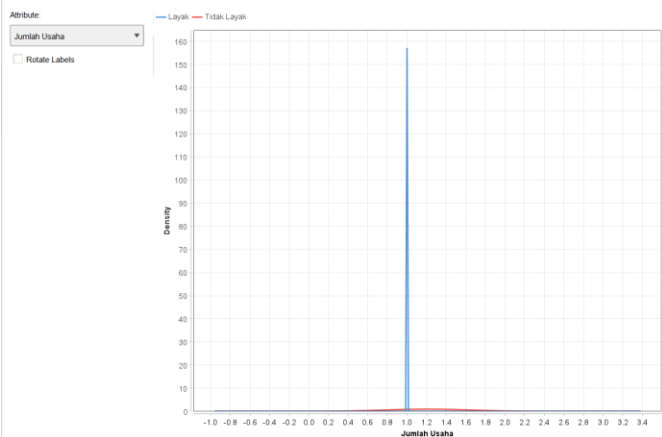
lebih layak untuk menerima bantuan dibandingkan dengan mereka yang memiliki pekerjaan.



Gbr. 23 Kelayakan Penerima BLT-UMKM Berdasarkan Pekerjaan

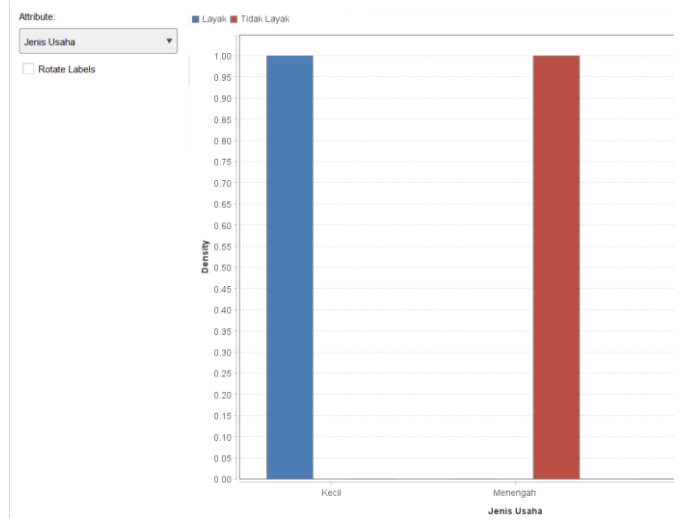
Pada Gambar 24 menunjukkan sebuah histogram distribusi kelayakan penerima bantuan langsung tunai bagi UMKM berdasarkan atribut “Jumlah Usaha”.

Dengan demikian, jumlah usaha yang mendominasi dalam dataset ini adalah 1, dan usaha dengan jumlah ini cenderung dianggap layak untuk menerima BLT-UMKM. Nilai-nilai jumlah usaha selain 1 sangat jarang muncul dalam data dan tidak menunjukkan kelayakan yang signifikan.



Gbr. 24 Kelayakan Penerima BLT-UMKM Berdasarkan Jumlah Usaha

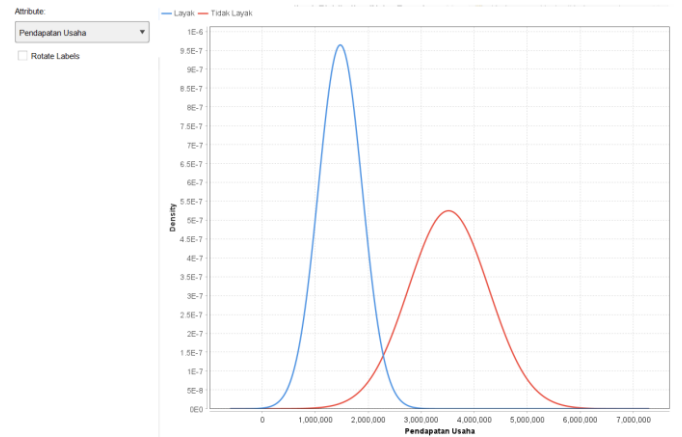
Pada Gambar 25 menunjukkan sebuah histogram distribusi kelayakan penerima bantuan langsung tunai bagi UMKM berdasarkan atribut “Jenis Usaha”.



Gbr. 25 Kelayakan Penerima BLT-UMKM Berdasarkan Jenis Usaha

Pada Gambar 26 menunjukkan sebuah histogram distribusi kelayakan penerima bantuan langsung tunai bagi UMKM berdasarkan atribut “Pendapatan Usaha”.

Dengan demikian, pendapatan usaha memiliki pengaruh yang signifikan terhadap kelayakan penerima BLT-UMKM. Penerima dengan pendapatan usaha lebih rendah cenderung lebih layak menerima bantuan, sedangkan penerima dengan pendapatan usaha lebih tinggi cenderung tidak layak menerima bantuan. Hal ini menunjukkan bahwa program BLT-UMKM ditargetkan untuk membantu usaha dengan pendapatan yang lebih rendah.



Gbr. 26 Kelayakan Penerima BLT-UMKM Berdasarkan Pendapatan Usaha

### E. Hasil Evaluasi Pengujian Naive Bayes

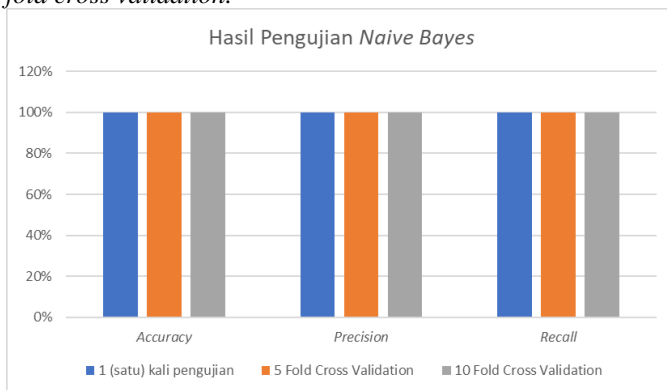
Setelah analisis klasifikasi *Naive Bayes* pada alat *RapidMiner* selesai, hasil klasifikasi *Naive Bayes* dibandingkan dengan 3 (tiga) opsi pengujian yang ditunjukkan pada Tabel 3. Dari hasil evaluasi didapatkan akurasi sebesar 100%, Presisi 100%, dan *Recall* 100%. Pada pengujian di atas dengan menggunakan *RapidMiner* terdapat 98 data yang diprediksi “Layak” dan 104 data yang diprediksi “Tidak Layak” dalam menerima BLT-UMKM.

TABEL III  
PREPARASI DATA

Model Evaluasi	Accuracy	Precision	Recall
1 (satu) kali pengujian	100%	100%	100%
5 Fold Cross Validation	100%	100%	100%
10 Fold Cross Validation	100%	100%	100%

Pada Gambar 27 menampilkan grafik hasil pengujian naïve bayes yang diukur berdasarkan 3 (tiga) opsi pengujian yaitu akurasi, presisi, dan *recall*. Pengujian dilakukan dengan tiga pendekatan berbeda.

Dengan demikian, algoritma *naïve bayes* menunjukkan performa yang sangat baik dan konsisten diketiga pendekatan pengujian (satu kali pengujian, 5 *fold cross validation* dan 10 *fold cross validation*).



Gbr. 27 Grafik Hasil Pengujian *Naïve Bayes*

#### IV. PENUTUP

##### A. Kesimpulan

Hasil penelitian dapat disimpulkan sebagai berikut:

- Hasil nilai probabilitas untuk kategori "Layak" adalah 0,485, dengan pembulatan nilai 0,48, dan nilai probabilitas untuk kategori "Tidak Layak" adalah 0,515, dengan pembulatan nilai 0,51.
- Melalui penelitian ini penulis mendapatkan *performance* prediksi "Layak" dan "Tidak Layak", akurasi yang didapatkan dengan 3 (tiga) opsi pengujian yaitu dengan 1 (satu) kali pengujian, 5 *fold cross validation* dan 10 *fold cross validation* sebesar 100%,

Presisi 100%, *Recall* 100%, dan AUC (*Area under curver*) 100%.

##### B. Saran

Adapun saran untuk penelitian lanjutan yaitu:

- Bagi penelitian selanjutnya diharapkan dapat membuat aplikasi untuk mengklasifikasi kelayakan penerima bantuan langsung tunai bagi UMKM khususnya di Kota Cimahi umumnya di seluruh Indonesia.
- Atribut kriteria yang digunakan sebagai penunjang dalam penentuan kelayakan penerima bantuan dapat lebih banyak seperti atribut jumlah pegawai, lokasi, pemodalan dan lainnya.
- Pengujian harus dilakukan dengan data yang sama tetapi menggunakan algoritma yang berbeda, seperti *Neural Network*, *Support Vector Machine*, C4.5 dan *K-Nearest Neighbor* untuk mendapatkan hasil yang lebih baik di masa mendatang.

#### REFERENSI

- Tambunan, T. Usaha Mikro Kecil, dan Menengah. Bogor: Ghalia Indonesia. 1998.
- Ilmiah, J., Erudisi, E. &, Yendra, M., Putri, W., & Wetsi, M. Dampak Bantuan Langsung Tunai (BLT) Dana Desa Covid-19 Terhadap Perekonomian Masyarakat. *JIEE: Jurnal Ilmiah Ekotrans & Erudisi*, 1(2), 14–22, 2021.
- Connolly, T., Begg, C., & Begg, C. Systems Database A Practical Approach to Design, Implementasi and anagement. *In Database Systems* (1–1425), 2010, www.booksites.net/connbegg.
- B. D. Laraswati. (2022) Algoritma Data Science School. [Online]. <http://blog.algoritma/data-mining/>
- Romli, I., & Zy, A. T. Penentuan Jadwal Overtime Dengan Klasifikasi Data Karyawan Menggunakan Algoritma C4.5. *Sains Komputer & Informatika (J-SAKTI)*, 4(2), 694–701, 2020.
- Gemasih, H., Nizar, I., & Rayuwati. Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid. *Jurnal Riset Rumpun Ilmu Teknik*, 1(1), 38–46, 2022.
- Supriyanto. Penerapan Algoritma Naive Bayes Classifier untuk Penentuan Bantuan Bedah Rumah Di Kelurahan Krapyak. Universitas Semarang: Teknik Informatika, 2020.
- Rahmadani, N., Risnawati, & Sena, M. D. Penerapan Algoritma Naive Bayes Dalam Penentuan Kelayakan Penerima Bantuan Program Keluarga Harapan. *Jurnal Teknologi Komputer dan Sistem Informasi*, 3(2), 40–48, 2023.
- Sugianto, C. A., & Maulana, F. R. Algoritma Naive Bayes untuk Klasifikasi Penerima Bantuan Pangan Non Tunai Di Kelurahan Utama. *Techno.COM*, 18(4), 321–331, 2019.
- Pebdika, A., Herdiana, R., & Solihudin, D. Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima PIP. *Jurnal Mahasiswa Teknik Informatika*, 7(1), 452–458, 2023.