

Penerapan Algoritma *K-Nearest Neighbor* Untuk Klasifikasi Kelayakan Status Penduduk Miskin Di Desa Susukan Tonggoh

Aas Nurjanah¹, Ahmad Rifai²

*Corresponding author : aasnurjanah0408@gmail.com

¹² STMIK IKMI CIREBON

Abstract

*Poverty is a condition where a person cannot meet his needs for a dignified life. Poverty in Indonesia does not only occur in one stage, but poverty extends from urban to rural areas. For example poverty in Susukan Tonggoh Village. In Susukan Tonggoh Village, the number of poor people continues to grow. The government actually plans various steps to overcome the problem of poverty in the village. However, the biggest problem is the lack of accurate information from the village about which residents qualify to be categorized as poor, making it difficult for the government's plans to address the problem of poverty. Therefore the purpose of this final project is to classify residents who are categorized as poor in Susukan Tonggoh Village. This final project was carried out based on data obtained from direct observation in Tonggoh Susukan Village as many as 282 data. The method used in this project is the data mining method, which is a classification method with the *K-Nearest Neighbor* algorithm using 6 regular attributes, namely name, address, occupation, place of residence, income level, source of income and 1 special attribute, namely the status of the poor. This final project creates qualification levels for villagers who meet poverty criteria, namely in the form of eligible and unfit categories. After running the entire process with a value of $K = 5$, this final project produces an accuracy rate of 95.11%. Predictions were not feasible and true were not feasible as many as 72. Predictions were not feasible and were feasible as many as 4. Predictions were feasible and were not feasible as many as 8. Predictions were feasible and true were feasible as many as 141. With a feasible class recall of 97.24% and class recall not feasible of 90.00 %.*

Keywords: Data Mining, K-Nearest Neighbor, Poverty, Classification.

Abstrak

Kemiskinan adalah suatu keadaan dimana seseorang tidak dapat memenuhi kebutuhannya akan kehidupan yang makmur. Kemiskinan di Indonesia tidak hanya terjadi dalam satu tahap, tetapi kemiskinan meluas dari perkotaan hingga pedesaan. Misalnya kemiskinan di Desa Susukan Tonggoh. Di Desa Susukan Tonggoh, jumlah penduduk miskin terus bertambah. Pemerintah sebenarnya merencanakan berbagai langkah untuk mengatasi masalah kemiskinan di desa. Namun, masalah terbesarnya adalah kurangnya informasi yang akurat dari desa tentang penduduk mana yang layak untuk dapat dikategorikan miskin, sehingga mempersulit rencana pemerintah untuk mengatasi masalah kemiskinan. Oleh karena itu tujuan dari penelitian ini adalah untuk mengklasifikasikan penduduk yang layak dikategorikan sebagai penduduk miskin di Desa Susukan Tonggoh. Penelitian ini dikerjakan berdasarkan data yang diperoleh dari pengamatan langsung di Desa Tonggoh

Susukan sebanyak 282 data. Metode yang digunakan dalam penelitian ini adalah metode *data mining*, yaitu metode klasifikasi dengan algoritma *K-Nearest Neighbor* menggunakan 6 atribut regular yaitu nama, alamat, pekerjaan, tempat tinggal, tingkat pendapatan, sumber pendapatan dan 1 atribut khusus yaitu status penduduk miskin. Penelitian ini menciptakan tingkat kualifikasi penduduk desa yang memenuhi kriteria kemiskinan, yaitu berupa kategori layak dan tidak layak. Setelah menjalankan seluruh proses dengan nilai $K=5$, penelitian ini menghasilkan tingkat akurasi sebesar 94.67 %. Prediksi tidak layak dan true tidak layak sebanyak 72. Prediksi tidak layak dan ternyata layak sebanyak 4. Prediksi layak dan ternyata tidak layak sebanyak 8. Prediksi layak dan true layak sebanyak 141. Dengan *class recall* layak sebesar 97.24% dan *class recall* tidak layak sebesar 90.00%.

Kata kunci: *Data Mining*, *K-Nearest Neighbor*, Kemiskinan, Klasifikasi.

Pendahuluan

Indonesia memiliki salah satu konsentrasi penduduk miskin tertinggi di dunia. (Nasution et al., 2020). Menurunnya jumlah penduduk miskin merupakan salah satu indikator keberhasilan pembangunan suatu negara. Menurut data Badan Pusat Statistik (BPS) tentang angka kemiskinan di Indonesia, jumlah penduduk miskin pada September 2020 sebanyak 27,55 juta jiwa atau 10,19% dari total penduduk Indonesia. Hal ini disebabkan oleh beberapa faktor antara lain pertumbuhan penduduk yang cepat, pengangguran yang tinggi, pertumbuhan ekonomi yang lambat, tingkat pendidikan yang rendah, dan distribusi yang tidak merata. (Rachma et al., 2022). Desa Susukan Tonggoh, salah satu desa di Kabupaten Cirebon, memiliki jumlah penduduk miskin yang cukup banyak dan Cirebon sendiri masuk dalam daftar 10 kabupaten/kota dengan penduduk miskin di Jawa Barat.

Meskipun pemerintah telah melakukan berbagai upaya dan program untuk mengentaskan kemiskinan di Desa Susukan Tonggoh, namun masalah tersebut tetap ada. Salah satu masalah besar yang sering menghambat keberhasilan program pemerintah adalah tidak tersedianya data sasaran yang tepat bagi masyarakat miskin. (Nurmayanti, 2021). Selain itu, ada penyebab lain, seperti banyaknya data status penduduk miskin di daerah yang tidak sesuai dengan kondisi nyata penduduk, seperti penduduk yang terdaftar di pemerintah desa sebagai penduduk miskin miskin tetapi tidak memenuhi syarat orang yang berhak digolongkan miskin dalam hal pendapatan, pekerjaan, dan kondisi kehidupan. Akibatnya, jumlah penduduk miskin di Desa Susukan Tonggoh harus diklasifikasi berdasarkan kelayakan status penduduk miskin. Mengingat semakin banyaknya data yang tersedia, maka pendekatan yang dilakukan dalam *data mining* adalah dengan menggunakan teknik klasifikasi. Dalam hal klasifikasi, algoritma *K-Nearest Neighbor* (*KNN*) harus digunakan. Algoritma *K-Nearest Neighbor* adalah metode klasifikasi yang menggunakan data pembelajaran yang paling dekat dengan data uji untuk mengklasifikasikan objek. (Badu, 2016).

Metode algoritma *K-Nearest Neighbor* bukanlah hal baru yang digunakan dalam klasifikasi, terbukti dari penelitian sebelumnya yang menggunakan teknik klasifikasi dengan algoritma *K-Nearest Neighbor*. Misalnya penelitian yang dilakukan oleh Kurnia, Fitra, Kurniawan, Jhoni, Fahmi, Ichsan, Monalisa, Siti

(2019) membahas tentang klasifikasi keluarga miskin dengan menggunakan metode *K-Nearest Neighbor*. (Kurnia et al., 2019).

Berdasarkan masalah yang sudah di paparkan, data tentang status penduduk miskin di Desa Susukan Tonggoh seringkali di tentukan secara tidak akurat. Sehingga diperlukan adanya klasifikasi kelayakan status penduduk miskin supaya adanya gambaran penduduk mana yang layak di kategorikan sebagai penduduk miskin agar dapat memudahkan pemerintah dalam menangani masalah kemiskinan secara merata. (Anuraga, 2015)

Metode Penelitian

2.1 Pengumpulan Data

Penelitian ini mengandalkan data primer. Penelitian ini mengandalkan data yang dikumpulkan melalui observasi langsung di Desa Susukan Tonggoh. Dengan total 282 data, data yang terkumpul adalah tentang penduduk miskin desa Susukan Tonggoh. Penulis mengumpulkan data pada tanggal 21 Januari 2023 pukul 10.00 WIB s.d. selesai.

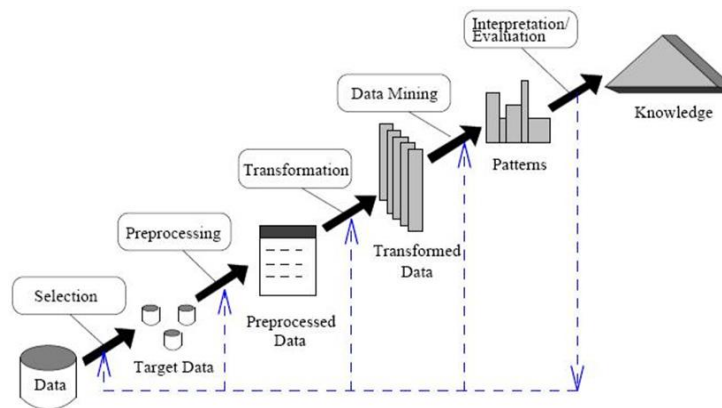
ID	NAMA	NIK	ALAMAT	PEKERJAAN	STATUS RUMAH	TINGKAT PENGHASILAN	SUMBER PENGHASILAN	TAMU	STATUS PENDUDUK MISKIN
1	ADHINAH	31090491030001	Blk Maria 002.001	Ibu rumah tangga	Rumah Sendiri	Sedang	Pendapatan suami	2022	Tidak layak
2	AEN NUR NISAH	310903661200007	Blk Kiwun 001.005	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan kerja	2022	Layak
3	AKHMAZ SOLIHAH	310903001020009	Blk Maria 001.001	Karyawan	Punya orang tua	Tinggi	Pendapatan kerja	2022	Tidak layak
4	AKHSA	31090491030001	Blk Kiwun 002.005	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan kerja	2022	Layak
5	ARAHM	310904204100004	Blk Pahang 002.001	Petani	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
6	ARHANTI	310904109500004	Blk Puhun 002.003	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
7	ARHANTI	310904109500004	Blk Kiwun 001.005	Buruh	Punya orang tua	Rendah	Pendapatan kerja	2022	Layak
8	ASMI SAMI	310904030400001	Blk Maria 001.001	Pejabat	Rumah Sendiri	Sedang	Pendapatan kerja	2022	Tidak layak
9	ASRIATI	310904040100005	Blk Village 002.004	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
10	ATI	310901295400008	Blk Pahang 002.001	Buruh	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
11	CASRIYAH	310904010400001	Blk Kiwun 002.005	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
12	CASRIYAH	310903001020009	Blk Pahang 002.002	Pejabat suami	Rumah Sendiri	Rendah	Pendapatan kerja	2022	Tidak layak
13	CHARDINO	310903005030003	Blk Maria 001.001	Buruh	Rumah orang tua	Rendah	Pendapatan kerja	2022	Layak
14	CHASMA	310903001020009	Blk Pahang 002.002	Buruh	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
15	DADINIAH	310903001020009	Blk Pahang 002.001	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
16	DANIRAH	310904911000001	Blk Kiwun 002.005	Ibu rumah tangga	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
17	DEVI	310904010400001	Blk Puhun 002.003	Karyawan	Sedang	Pendapatan suami	2022	Layak	
18	DEVI ERMAHATI	310904040100005	Blk Kiwun 002.001	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
19	DIO SUNDARI	310904020000001	Blk Maria 001.001	Petani	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
20	DIWA SUNDARI	310901101010005	Blk Puhun 002.003	Buruh	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
21	ENDIK SUDINI	310903001020009	Blk Maria 001.001	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
22	ERIKAWATIEM	310904020000001	Blk Maria 001.001	Karyawan	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
23	ERIKAWATI	310903001020009	Blk Maria 001.001	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
24	ERMAHATI	310903001020009	Blk Puhun 002.003	Petani	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
25	ETI RIZAHATI	310904020000001	Blk Maria 001.001	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
26	FATHIA FATHIAHATI	310903001020009	Blk Pahang 001.002	Petani	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Layak
27	HADIYAH	310903001020009	Blk Puhun 002.003	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
28	HANI	310903001020009	Blk Pahang 002.001	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
29	HANIRAH	310903001020009	Blk Kiwun 002.005	Petani	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
30	HANIT	310904010400001	Blk Kiwun 002.005	Ibu rumah tangga	Karyawan	Sedang	Pendapatan suami	2022	Layak
31	HANINDAHATI	310904010400001	Blk Kiwun 001.005	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
32	HA TARBIYAH	310903001020009	Blk Kiwun 001.005	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
33	HA WARDANI	310903001020009	Blk Maria 001.001	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
34	HAZI	310904102000003	Blk Kiwun 002.005	Wiraswasta	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
35	HAZI NUR HAHI	310903001020009	Blk Village 002.004	Wiraswasta	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
36	HAZIL	310903001020009	Blk Village 001.004	Wiraswasta	Rumah Sendiri	Tinggi	Pendapatan kerja	2022	Tidak layak
37	HAZIL	310904040100005	Blk Pahang 002.001	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
38	HANILAH	310903001020009	Blk Kiwun 001.005	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak
39	HEI	310904010400001	Blk Kiwun 001.005	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan suami	2022	Layak

Gambar 2.1 dataset penduduk miskin di desa susukan tonggoh

Metode pengumpulan data penelitian ini adalah observasi. Observasi adalah teknik pengumpulan data kualitatif yang melibatkan pengamatan langsung di lapangan atau lingkungan penelitian. Penulis melakukan observasi di desa Susukan Tonggoh, Kecamatan Susukan Lebak, Kabupaten Cirebon.

2.2 Tahapan Penelitian

Metode penelitian yang digunakan penulis dalam Menerapkan Algoritma *K-Nearest Neighbor* untuk klasifikasi kelayakan status penduduk miskin di Desa Susukan Tonggoh adalah metode dengan proses *Knowledge Discovery in Databases (KDD)*: Berikut tahapan-tahapan yang akan dilakukan dalam penelitian ini :



Gambar 2.2 Alur Penelitian

1. Seleksi Data (*Selection*)

Pada tahap ini di seleksi atribut mana yang akan dipakai dalam pemrosesan data mining menggunakan *select atribut*. Dari 9 atribut yaitu id, nik, nama penduduk, alamat, pekerjaan, status rumah, tingkat penghasilan, sumber penghasilan dan tahun akan di seleksi menjadi 6 atribut regular yang di butuhkan yaitu nama penduduk, alamat, pekerjaan, status rumah, tingkat penghasilan, sumber penghasilan, dan 1 atribut spesial yaitu status penduduk miskin.

2. Pemilihan Data (*Preprocessing/Cleaning*)

Preprocessing adalah proses menghilangkan data duplikat, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan data seperti kesalahan ketik dan data yang hilang. Operator *replace missing values*. akan digunakan untuk menghapus kumpulan data tanpa nilai (hilang). (Fadlan et al., 2018).

3. Transformasi (*Transformation*)

Pada tahap ini yang dilakukan adalah menkonversi entuk data yang belum mempunyai entitas dengan jelas pada bentuk data yang valid atau siap untuk dilakukan proses *Data Mining*. (Ali et al., 2021)

4. *Data Mining*

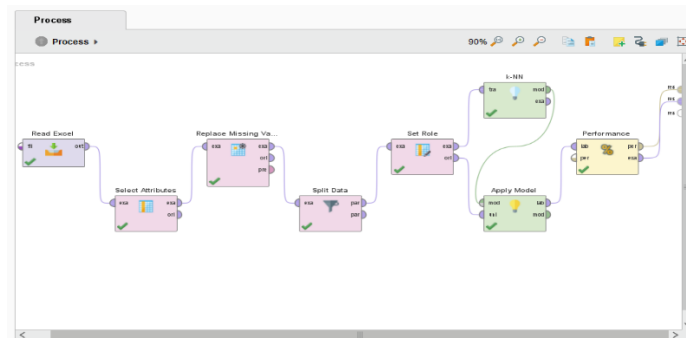
Pada tahap ini data akan di proses menggunakan algoritma *K-Nearest Neighbord*. Proses data mining terhadap status penduduk miskin Desa Susukan Tonggoh akan menggunakan *tools Rapid Miner*..

5. Interpretasi/Evaluasi (*Interpretation / Evaluation*)

Pada tahap ini, tahap interpretasi/evaluasi adalah proses menghasilkan output yang dapat dipahami dari proses pola informasi data mining. (Ramadani et al., 2019)

Hasil dan Pembahasan

Klasifikasi kelayakan status penduduk miskin di Susukan Tonggoh dilakukan dengan menggunakan aplikasi *Rapidminer*.

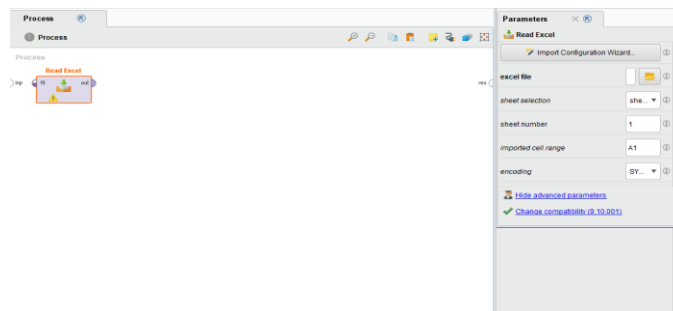


Gambar 3. 1 Proses Klasifikasi Menggunakan *Rapidminer*

Berdasarkan Gambar 4.1, proses klasifikasi menggunakan beberapa operator yang terdapat pada *Rapidminer*. Operator yang digunakan tersebut memiliki fungsi dan tujuan masing – masing serta berada pada tahapan yang berbeda. Tahapan pada proses klasifikasi adalah sebagai berikut :

1. Data Selection

Proses pertama dilakukan pada pemilihan data yaitu memasukkan operator “*read excel*” yang berfungsi untuk membaca data dari *file excel* penduduk miskin desa Susukan Tonggoh. Langkah selanjutnya adalah mengimpor data.



Gambar 3. 2 memasukkan operator *read excel*

Pada tahap impor data, sistem akan menampilkan *preview* data sebelum memasukkannya. Tujuannya adalah untuk memeriksa kembali apakah informasi yang dimasukkan sudah benar. Kemudian klik *Finish* untuk menyelesaikan proses pemilihan atribut.

Import Data - Format your columns.

Format your columns.

Replace errors with missing values

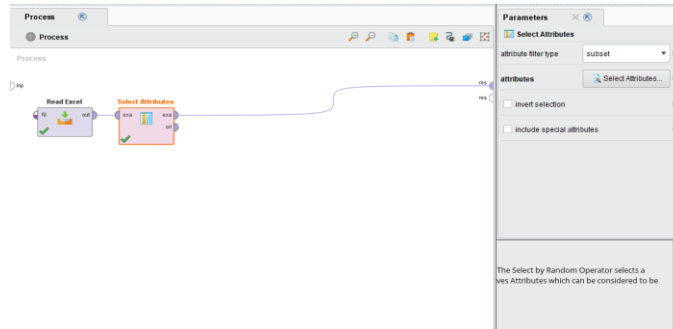
ID	NAMA	NIK	ALAMAT	PEKERJAAN	STATUS R...
Integer	polynomial	real	polynomial	polynomial	polynomial
1	ADMIRAH	3209084902830...	Blok Manis 002/...	Ibu rumah tangga	Rumah Sendiri
2	AEN NURANAH	3209086611900...	Blok Kliwon 001/...	Ibu rumah tangga	Punya orang tua
3	AHMAD BOLKIAH	3209081001950...	Blok Manis 001/...	Karyawan	Punya orang tua
4	AISEYAH	3209085402600...	Blok Kliwon 003/...	Ibu rumah tangga	Rumah Sendiri
5	ARMINI	3209084204710...	Blok Pahing 002/...	Petani	Rumah Sendiri
6	ARYANTI	3209084109500...	Blok Puhun 002/...	Ibu rumah tangga	Punya orang tua
7	ARYANTO	3209084310880...	Blok Kliwon 001/...	Buruh	Punya orang tua
8	ASMI SAM	3209084308430...	Blok Manis 001/0...	Pedagang	Rumah Sendiri
9	ASWITI	3209084604810...	Blok Wage 002/0...	Ibu rumah tangga	Punya orang tua
10	ATA	3209081504540...	Blok Pahing 002/...	Buruh	Rumah Sendiri
11	CASMINAH	3209084703660...	Blok Kliwon 003/...	Ibu rumah tangga	Punya orang tua
12	CASMIHAH	3209086106830...	Blok Pahing 002/...	Datamak acam	Rumah Sendiri

no problems.

Previous Finish Cancel

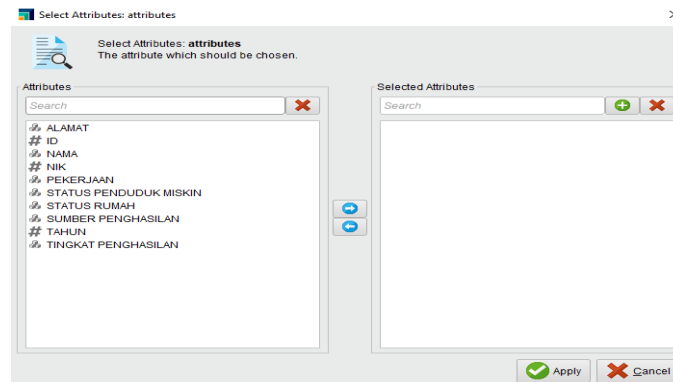
Gambar 3. 3 proses *import* data

Proses selanjutnya yaitu memasukan operator *select atribut*. Fungsi operator ini untuk membedakan mana atribut yang tidak digunakan dan mana yang digunakan dalam proses data mining.(Astuti & Guntara, 2018).



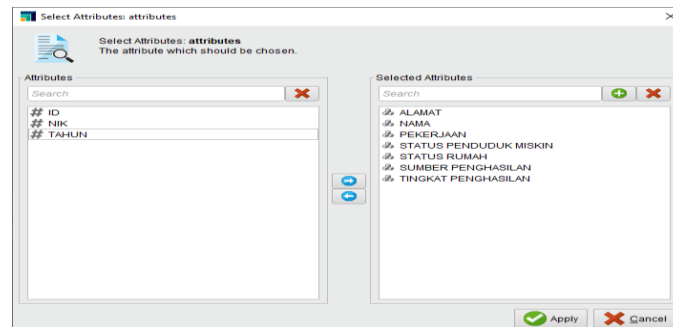
Gambar 3. 4 memasukan operator *select atribut*

Proses *select atribut* dapat dilihat pada gambar 4.5 atribut yang digunakan yaitu nama penduduk, alamat, pekerjaan, status rumah, tingkat penghasilan, sumber penghasilan, dan status penduduk miskin akan dipisahkan kedalam parameter *select atributs*



Gambar 3. 5 proses *select atribut*

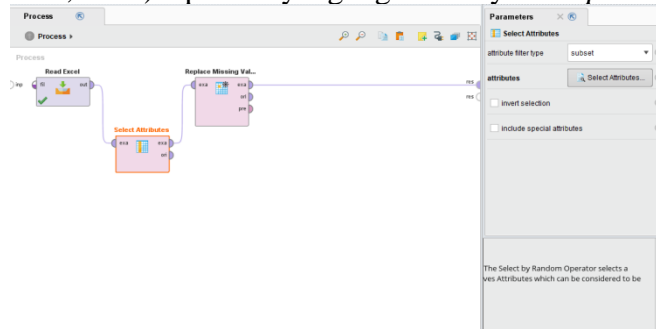
Hasilnya adalah dari 10 atribut, hanya menyisakan 7 atribut saja yang akan di proses dalam *data mining*.



Gambar 3. 6 hasil proses *select atribut*

2. Data preprocessing/cleaning

Dengan kemungkinan adanya data yang tidak relevan, proses *cleaning* tentu diperlukan. Data yang kosong, ganda dan missing akan dihilangkan, sebab data yang tidak relevan akan mengurangi tingkat akurasi pada pemrosesan *data mining*. (Putri et al., 2021) Operator yang digunakan yaitu *replace missing values*.



Gambar 3. 7 memasukan operator *replace missing values*

Dengan operator *replace missing values* akan diketahui data mana yang perlu diperbaiki seperti data yang kosong atau ganda.

	NAMA	ALAMAT	PEKERJAAN	STATUS R...	TINGKAT P...	SUMBER P...
	polynomial	polynomial	polynomial	polynomial	polynomial	polynomial
9	ASWITI	Blok Wage 002/0...	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan sua...
10	ATA	Blok Pahing 002/...	Buruh	Rumah Sendiri	Tinggi	Pendapatan kerja
11	CASMINAH	Blok Kliwon 003/...	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan sua...
12	CASMIH	Blok Pahing 002/...	Peternak ayam	Rumah Sendiri	?	Pendapatan kerja
13	DARSONO	Blok Manis 001/0...	Buruh	Rumah orang tua	Rendah	Pendapatan kerja
14	DASMA	Blok Pahing 002/...	Buruh	Rumah Sendiri	Tinggi	Pendapatan kerja
15	DASMINAH	Blok Pahing 002/...	Ibu rumah tangga	Punya orang tua	?	Pendapatan sua...
16	DASMINAH	Blok Kliwon 003/...	?	Rumah Sendiri	Tinggi	Pendapatan kerja
17	DATI	Blok Puhun 002/...	Ibu rumah tangga	Kontrakan	Sedang	Pendapatan sua...
18	DEWI ERNAWATI	Blok Kliwon 005/...	Ibu rumah tangga	Punya orang tua	Rendah	Pendapatan sua...
19	DIDI SUNARDI	Blok Manis 002/0...	Petani	Rumah Sendiri	Tinggi	Pendapatan kerja
20	EMAH SUHERMAN	Blok Puhun 001/...	Buruh	Rumah Sendiri	Tinggi	Pendapatan kerja

Gambar 3.8 proses *replace missing values*

Data yang digunakan setelah proses *cleaning* dengan operator *replace missing values* adalah data yang lengkap, artinya tidak ada data yang hilang. (Alfiansyah et al., 2022). Sehingga, proses transformasi data dapat dimulai.

Name	Type	Missing	Statistics	Filter (7 / 7 attributes)	Search for Attributes
✓ NAMA	Polynomial	0	Least: YULI HIDRAWATI (1) Most: DASMINAH (2) Values: DASMINAH (2), UAH HADIJAH (2)		
✓ ALAMAT	Polynomial	0	Least: Blok Kliwon 005/001 (1) Most: Blok Pahing 002/002 (33) Values: Blok Pahing 002/002 (33), Blok		
✓ PEKERJAAN	Polynomial	0	Least: Peternak bebek (1) Most: Ibu rumah tangga (120) Values: Ibu rumah tangga (120), Buruh		
✓ STATUS RUMAH	Polynomial	0	Least: Rumah orang tua (1) Most: Punya orang tua (117) Values: Punya orang tua (117), Kontrak		
✓ TINGKAT PENGHASILAN	Polynomial	0	Least: Pendapatan kerja (2) Most: Rendah (121) Values: Rendah (121), Tinggi (89), ... [2		
✓ SUMBER PENGHASILAN	Polynomial	0	Least: Pendapatan suami (117) Most: Pendapatan kerja (164) Values: Pendapatan kerja (164), Penda		
✓ STATUS PENDUDUK MISKIN	Polynomial	0	Least: Tidak layak (100) Most: Layak (181) Values: Layak (181), Tidak layak (100)		

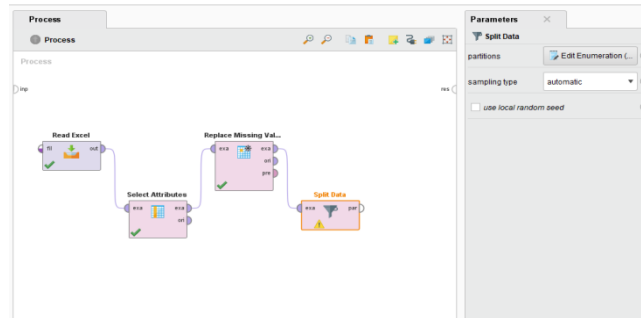
Gambar 3.9 hasil *replace missing values*

3. Data transformation

Proses transformasi ini melibatkan perubahan data yang dipilih menjadi format yang dapat diproses dalam *data mining*. Pada tugas akhir ini dilakukan dua proses transformasi data yaitu:

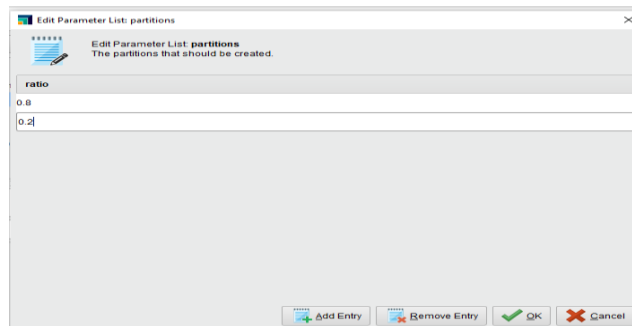
a. Proses pertama adalah membagi data (*split data*)

Pada klasifikasi *split data* operator digunakan untuk membagi data menjadi data *training* dan data *testing*, yang kemudian dibandingkan. (Prihandari, 2022).



Gambar 3. 10 proses *split data*

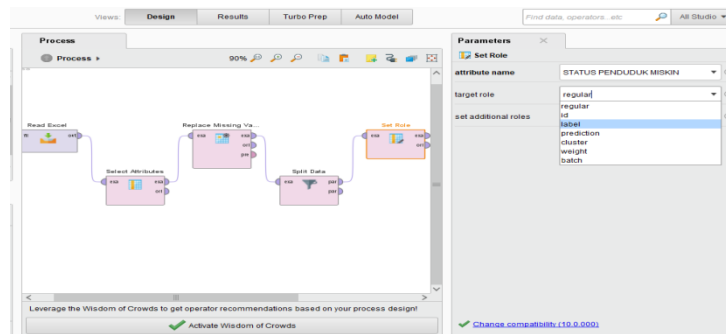
Pembagian data untuk dataset penentuan kelayakan status penduduk miskin yang berjumlah 282 data ini dilakukan dengan membagi data sebesar 80% untuk data *training* dan 20% untuk data *testing*.



Gambar 3. 11 parameter *split data*

b. Proses *transformation* yang kedua yaitu *set role*

Dalam proses *data mining* klasifikasi, diperlukan sebuah atribut spesial yang dijadikan sebagai label. Atribut label pada klasifikasi berfungsi sebagai atribut penentu, nilai dari atribut label akan menjadi hasil akhir dari klasifikasi. (Desa et al., 2019). Operator *set role* digunakan untuk merubah atribut *regular* menjadi atribut spesial, dalam hal ini yaitu label. Pada proyek tugas akhir ini, atribut yang menjadi penentu adalah atribut status penduduk miskin, karenanya atribut status penduduk miskin harus diubah *rolenya* dari yang semula merupakan *regular attribute* menjadi *special attribute* yaitu label dengan menggunakan operator *Set Role*.



Gambar 3. 12 proses set role

Row No.	STATUS PE.	NAMA	ALAMAT	PEKERJAAN	STATUS ROL.	TINGKAT PE.	SUMBER PE.
1	Tidak layak	ACMERAH	Blok Manis 0.	Ibu rumah ta.	Rumah Sendiri	Sedang	Pendapatan ...
2	Layak	AENHURAN	Blok Kiliwon	Ibu rumah ta.	Punya orang L.	Rendah	Pendapatan ...
3	Tidak layak	AHMAD BOL.	Blok Manis 0.	Karyawan	Punya orang L.	Tinggi	Pendapatan ...
4	Layak	ADYAH	Blok Kiliwon	Ibu rumah ta.	Punya orang L.	Rendah	Pendapatan ...
5	Tidak layak	ARMAN	Blok Pating	Petani	Rumah Sendiri	Tinggi	Pendapatan ...
6	Layak	ARYANTI	Blok Puhun 0.	Ibu rumah ta.	Punya orang L.	Rendah	Pendapatan ...
7	Layak	ARYANTO	Blok Kiliwon 0.	Buruh	Punya orang L.	Rendah	Pendapatan ...
8	Tidak layak	ASMI SAM	Blok Manis 0.	Pedagang	Rumah Sendiri	Sedang	Pendapatan ...
9	Layak	ASWITI	Blok Wage 0.	Ibu rumah ta.	Punya orang L.	Rendah	Pendapatan ...
10	Tidak layak	ATA	Blok Pating	Buruh	Rumah Sendiri	Tinggi	Pendapatan ...
11	Layak	CASIRAH	Blok Kiliwon 0.	Ibu rumah ta.	Punya orang L.	Rendah	Pendapatan ...
12	Tidak layak	CASIRAH	Blok Pating	Peternak ayam	Rumah Sendiri	Rendah	Pendapatan ...
13	Layak	DARSONO	Blok Manis 0.	Buruh	Rumah oran.	Rendah	Pendapatan ...
14	Tidak layak	DASMA	Blok Pating	Buruh	Rumah Sendiri	Tinggi	Pendapatan ...
15	Layak	DASIRAH	Blok Pating	Ibu rumah ta.	Punya orang L.	Rendah	Pendapatan ...

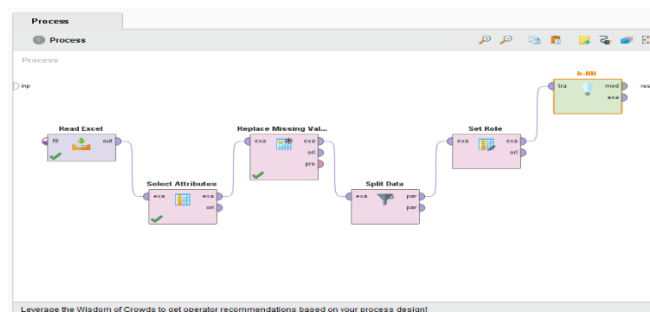
Gambar 3. 13 hasil proses set role

4. Data mining

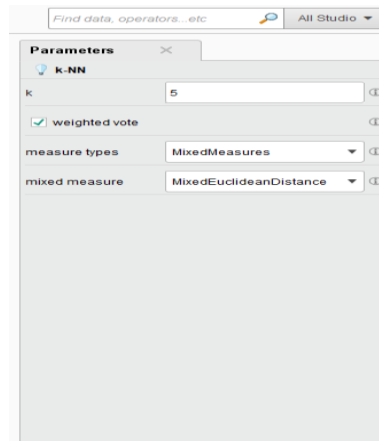
Langkah-langkah klasifikasi dengan menerapkan algoritma *K-Nearest Neighbor* ini dirancang dengan menggunakan tools *Rapidminer*. Pada proyek tugas akhir ini, digunakan beberapa operator yang ada pada *Rapidminer*, yaitu :

a. *K-Nearest Neighbord* (KNN)

Algoritma yang dipakai untuk proyek tugas akhir ini yaitu algoritma *K-Nearest Neighbord* (KNN). Setelah dataset sudah diolah pada tahap sebelumnya, kemudian dataset akan diproses dengan menggunakan operator KNN pada *Rapidminer*.



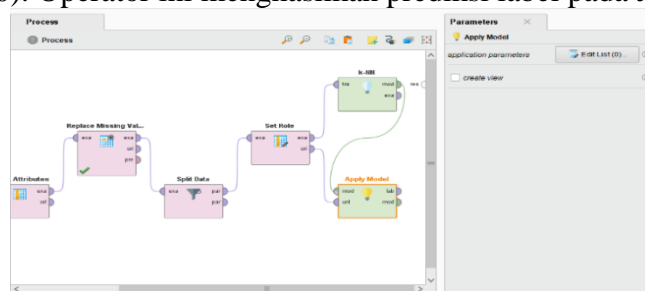
Gambar 3. 14 proses *K-Nearest Neighbord* (KNN)



Gambar 3. 15 parameter *K-Nearest Neighbor (KNN)*

b. *Apply Model*

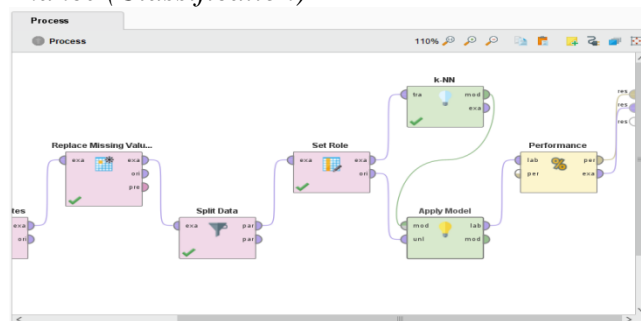
Operator *Apply Model* adalah operator lanjutan dari *K-Nearest Neighbor* yang digunakan untuk mengaplikasikan model yang sebelumnya telah dilatih menggunakan *data training* pada *data testing*. (Rahmawati et al., 2020). Operator ini menghasilkan prediksi label pada *data testing*.



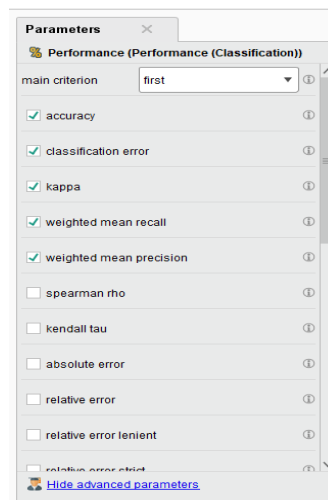
Gambar 3. 16 proses *apply model*

Setiap data mempunyai nilai *confidence* yang berbeda pada masing-masing *value* labelnya. (Annur, 2018). Berdasarkan Gambar 4.6 dapat disimpulkan bahwa *value* label yang mempunyai nilai *confidence* lebih tinggi akan menjadi *value* yang terpilih sebagai label prediksi. Berikut adalah sampel data *testing* yang telah diproses dengan operator *Apply Model* :

c. *Performance (Classification)*



Gambar 3.17 proses *performance*



Gambar 3. 18 parameter *performance (classification)*

5. Interpretasi/Evaluasi

Nilai Akurasi :

	true Tidak layak	true Layak	class precision
pred. Tidak layak	72	4	94.74%
pred. Layak	8	141	94.63%
class recall	90.00%	97.24%	

Gambar 3. 19 hasil nilai akurasi

Hasil Penelitian :

Penelitian tentang klasifikasi kelayakan status penduduk miskin desa Susukan Tonggoh dengan menggunakan algoritma *K-Nearest Neighbord*, menciptakan hasil dengan rincian sebagai berikut :

- a. Nilai akurasi sebesar 94.67%
- b. Prediksi tidak layak dan true tidak layak sebanyak 72.
- c. Prediksi tidak layak dan ternyata layak sebanyak 4.
- d. Prediksi layak dan ternyata tidak layak sebanyak 8.
- e. Prediksi layak dan true layak sebanyak 141.
- f. Dengan *class recall* layak sebesar 97.24% dan *class recall* tidak layak sebesar 90.00%

Kesimpulan

Berdasarkan penelitian yang telah dilakukan, diperoleh kesimpulan sebagai berikut :

1. Hasil yang didapat pada penelitian dan pembahasan dengan melakukan klasifikasi dengan menggunakan metode algoritma *K-Nearest Neighbor* ini menghasilkan nilai akurasi yang baik. Penelitian yang dilakukan menggunakan metode klasifikasi pada *rapidminer* dengan menggunakan algoritma *K-Nearest Neighbor* serta didukung oleh data penduduk miskin pada Desa Susukan Tonggoh tahun 2022. Penelitian ini menghasilkan tingkat akurasi, hasil *precision* dan hasil *recall*, yang mengacu pada status penduduk miskin yang dijadikan label. Hasil akurasi dari penelitian ini menghasilkan tingkat akurasi sebesar 94.67%. Hasil *precision* Layak sebesar 94.74%, Tidak Layak sebesar 94.63%. Hasil dari *recall* sebesar 97.24% untuk hasil dari Layak dan 90.00% untuk hasil dari Tidak Layak.
2. Berdasarkan hasil pengujian menggunakan tools Rapidminer menghasilkan nilai evaluasi *performance* yang optimal berada di $k = 5$ memiliki nilai akurasi 94.67% dengan 7 atribut data sebagai variabel yang berpengaruh dan ditentukan oleh aspek pekerjaan, status rumah, tingkat penghasilan, dan sumber penghasilan.

Daftar Pustaka

- Alfiansyah, D. N., Nastiti, V. R. S., & Hayatin, N. (2022). Penerapan Metode K-Means pada Data Penduduk Miskin Per Kecamatan Kabupaten Blitar. *Jurnal Repositor*, 4(1), 49–58. <https://doi.org/10.22219/repositor.v4i1.1416>
- Ali, I., Ade Kurnia, D., Pratama, M. A., & Al Ma'ruf, F. (2021). KOPERTIP: Jurnal Ilmiah Manajemen Informatika dan Komputer Klasifikasi Status Stunting Balita Di Desa Slangit Menggunakan Metode K-Nearest Neighbor. *Jurnal Ilmiah Manajemen Informatika Dan Komputer*, 05(03), 35–38. <http://jurnal.kopertipindonesia.or.id/index.php/kopertip>
- Annur, H. (2018). Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes. *ILKOM Jurnal Ilmiah*, 10(2), 160–165. <https://doi.org/10.33096/ilkom.v10i2.303.160-165>
- Anuraga, G. (2015). Hierarchical Clustering Multiscale Bootstrap untuk Pengelompokkan Kemiskinan di Jawa Timur. *Statistika*, 1, 27–33.
- Astuti, F. D., & Guntara, M. (2018). Analisis Performa Algoritma K-NN Dan C4.5 Pada Klasifikasi Data Penduduk Miskin. *Jurnal Rekayasa Teknologi Informasi (JURTI)*, 2(2), 135. <https://doi.org/10.30872/jurti.v2i2.1865>
- Badu, Z. S. (2016). *PENERAPAN ALGORITMA K-NEAREST NEIGHBOR*. November.
- Desa, D. I., Kidul, N., Pati, K., & Tengah, J. (2019). *K-nearest neighbor*. 4(1), 21–28.

- Fadlan, C., Ningsih, S., & Windarto, A. P. (2018). Penerapan Metode Naïve Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra. *Jurnal Teknik Informatika Musirawas (JUTIM)*, 3(1), 1. <https://doi.org/10.32767/jutim.v3i1.286>
- Kurnia, F., Kurniawan, J., Fahmi, I., & Monalisa, S. (2019). *Klasifikasi Keluarga Miskin Menggunakan Metode K- Nearest Neighbor Berbasis Euclidean Distance*. November, 230–239.
- Nasution, I., Windarto, A. P., & Fauzan, M. (2020). Penerapan Algoritma K-Means Dalam Pengelompokan Data Penduduk Miskin Menurut Provinsi. *Building of Informatics, Technology and Science (BITS)*, 2(2), 76–83. <https://doi.org/10.47065/bits.v2i2.492>
- Nurmayanti, W. P. (2021). Penerapan Naive Bayes dalam Mengklasifikasikan Masyarakat Miskin di Desa Lepak. *Geodika: Jurnal Kajian Ilmu Dan Pendidikan Geografi*, 5(1), 123–132. <https://doi.org/10.29408/geodika.v5i1.3430>
- Putri, A. C., Hariyanto, F. E., Andini, N. L. E., & Zulkarnaen, Z. C. S. (2021). Klasifikasi Rumah Tangga Miskin di Provinsi Papua Tahun 2017 Menggunakan Metode Naive Bayes. *Jurnal Sains Matematika Dan Statistika*, 7(1), 89. <https://doi.org/10.24014/jsms.v7i1.11924>
- Rachma, C. A., Matematika, P. S., Sains, F., Teknologi, D. A. N., Islam, U., Maulana, N., & Ibrahim, M. (2022). *Implementasi algoritma k-nearest neighbor dalam penentuan klasifikasi tingkat kedalaman kemiskinan provinsi jawa timur*.
- Rahmawati, D., Studi, P., Informatika, T., Sains, F., Teknologi, D. A. N., Islam, U., Syarif, N., & Jakarta, H. (2020). *BANTUAN SOSIAL DENGAN MENGIMPLEMENTASIKAN ALGORITMA K-NEAREST NEIGHBOR (Studi Kasus : RW 13 Kelurahan Palmerah Jakarta Barat) BANTUAN SOSIAL DENGAN MENGIMPLEMENTASIKAN ALGORITMA K-NEAREST NEIGHBOR (Studi Kasus : RW 13 Kelurahan Palmerah Jakarta Barat)*.
- Ramadani, S., Ambarita, I., & Pardede, A. M. H. (2019). Metode K-Means Untuk Pengelompokan Masyarakat Miskin Dengan Menggunakan Jarak Kedekatan Manhattan City Dan Euclidean (Studi Kasus Kota Binjai). *Information System Development (ISD)*, 04(2), 15–29.