ANALISIS TINGKAT KEMISKINAN DENGAN MENGGUNAKAN *ROUGH SET*

RIKE MARLIANI, S.SI

Badan Pusat Statistik Kota Samarinda Provinsi Kalimantan Timur

rike@bps.go.id

**ABSTRAK**

Kemiskinan merupakan isu yang sangat penting dan sering dijadikan acuan terhadap kesuksesan pembangunan. Terdapat banyak upaya yang telah dilakukan untuk mengentaskan pemerintah seperti mengurangi pengangguran, meningkatkan sumber daya masyarakat, menjaga inflasi dan menggenjot perekonomian baik daerah secara khususnya ataupun nasional pada umumnya. Penelitian ini menggunakan model berbasis *rough set* untuk mengetahui faktor yang mempengaruhi kemiskinan di Provinsi Kalimantan Timur. Penelitian ini menggunakan data dari BPS Provinsi Kalimantan Timur dari tahun 2006-2020. Dari hasil penelitian menunjukan bahwa tingkat pengangguran merupakan faktor utama yang mempengaruhi kemiskinan diprovinsi Kalimantan Timur.

Kata Kunci: *Rough set*, kemiskinan, pengangguran, *Rule Extraction*, Kalimantan Timur.

***ABSTRACT***

*Poverty is a very important issue and is often used as a reference for successful development. There have been many efforts that have been made to alleviate the government, such as reducing unemployment, increasing public resources, maintaining inflation and boosting the economy both in the region in particular and nationally in general. This study uses a rough set-based model to determine the factors that influence poverty in East Kalimantan Province. This study uses data from the BPS of East Kalimantan Province from 2006-2020. The results of the study indicate that the unemployment rate is a major factor affecting poverty in the province of East Kalimantan.*

Kata Kunci: *Rough set*, *poverty, unemployment, rule extraction*, *East Kalimantan*.

1. Pendahuluan

Indonesia merupakan negara yang memiliki jumlah penduduk paling banyak keempat di dunia dari laporan Departemen Populasi Divisi Urusan Sosial dan Ekonomi Perserikatan Bangsa-Bangsa (PBB) pada Juni 2017. Sebagai negara yang belum seabad merdeka, masalah kependudukan menjadi masalah yang besar bagi pemerintah. Tingkat kemiskinan dan kesejahteraan masyarakat menjadi masalah utama yang hingga saat ini masih sangat berpengaruh bagi kemajuan negara.

Kemiskinan merupakan isu yang sangat penting karena terkait dengan kesejahteraan penduduk yang tinggal pada daerah tersebut. Selain itu, kemiskinan umumnya menjadi faktor penyebab dari tindak kejahatan dan kriminalitas. Kemiskinan diartikan sebagai suatu keadaan dimana seseorang tidak sanggup memelihara dirinya sendiri sesuai dengan taraf kehidupan kelompok dan juga tidak mampu memanfaatkan tenaga mental, maupun fisiknya dalam kelompok tersebut.

Dalam Laporan Badan Pusat Statistik (BPS) bulan Februari 2021 tentang tingkat kemskinan di Provinsi Kalimantan Timur diketahui bahwa terjadi peningkatan persentase penduduk penduduk miskin di perkotaan dari 4,45 persen menjadi 5,10 persen sementara untuk penduduk miskin di pedesaan juga mengalami peningkatan dari 9,51 persen menjadi 9,98 persen. Meningkatnya persentasi kemiskinan di Provinsi Kalimantan Timur membuat penulis ingin menganalisa faktor-faktor yang mempengaruhi kemiskinan di Provinsi Kalimantan Timur sehingga pemerintah dapat fokus pada perbaikan faktor-faktor tersebut dan dapat menurunkan tingkat kemiskinan.

Berdasarkan penelusuran literatur, banyak penelitian terkait dengan prediksi kemiskinan. Beberapa metode yang digunakan untuk menganalisa seperti jaringan saraf tiruan [1], K-Mean [2], pohon keputusan (C4.5 dan C4.5 berbasis *back propagation*) [3], *back propagation* [4], *moving average*, naïve bayes, regresi linier berganda[5] dan rough set[6]. Namun, hanya sedikit literatur yang membahas mengenai rough set dan kemiskinan. Rough set merupakan *mathematic tool* yang sangat handal, ia dapat menganai data yang *vague, incomplete* dan *inconsistent*. Selain itu, kelebihan lain dari metode ini ialah tidak memerlukan parameter tambahan karena informasi seperti probabilistik diperoleh dari data itu sendiri [7]. Hal ini berbeda dengan teknik yang digunakan pada penelitian sebelumnya yang memerlukan parameter tambahan yang perlu ditentukan seperti K-*mean* harus mencari nilai $K$ yang optimum atau jaringan syaraf tiruan yang perlu menentukan jumlah layar dan perseptron yang digunakan untuk menghasilkan performa yang baik. Selain itu metode ini dapat menangani data yang tidak lengkap (*incomplete information system*)[8]. Berdasarkan hal tersebut maka penelitian ini akan menggunakan rough set untuk menganalisa tingkat kemiskinan pada provinsi Kalimantan Timur.

1. Tinjauan Pustaka
	1. Kemiskinan

Di Indonesia pengukuran kemiskinan menggunakan kriteria dari BPS. Menurut BPS kriteria kemiskinan menggunakan pendekatan kebutuhan dasar (*basic needs*). Untuk perhitungan *Headcount index* sebagai dasar garis kemiskinan BPS ditentukan berdasarkan batas pengeluaran minimum untuk konsumsi makanan setara dengan 2100 kalori per hari dan konsumsi nonmakanan.

Isu kemiskinan sangat penting, sehingga banyak penelitian yang berusaha untuk mempredikasi tingkat kemiskinan seperti Wanto (2018) memprediksi jumlah kemiskinan dengan menggunakan jaringan syaraf tiruan (JST) pada kabupaten di provinsi Riau. Mamase dan Sinukun [9] melakukan prediksi tingkat kemiskinan pada provinsi Gorontalo dengan menggunakan metode GRNN (*Generalized Regression Neural Network*) dengan MAPE (*Mean Absolute Persentage Error*) digunakan sebagai kriteria evaluasi model prediksi dengan nilai MAPE sebesar 0.03 memberikan prediksi tingkat kemiskinan yang cukup akurat. Kemudian untuk mengatasi masalah *overfitting*, Mamase dan Sinukun [2] menggunakan K-Mean untuk pemilihan data latih yang sesuai sehingga menghasilkan nilai MAPE 2.3% (0.023). Penggunaan jaringan saraf tiruan (*neural network*) memberikan hasil yang baik, namun teknik ini memiliki kesulitan dalam hal pemilihan jumlah layar dan perseptron pada jaringan untuk menghasilkan performa yang baik.

Rahmadi, dkk [10] membuat model klasifikasi dengan menggunakan Decision Tree (Pohon keputusan) C5.0. Penggunaan Decision Tree tidak hanya digunakan untuk melakukan klasifikasi, namun juga digunakan untuk mengungkap faktor-faktor yang mempengaruhinya dari berbagai banyak faktor yang ada. Pada penelitian ini atribut yang digunakan ialah keluhan kesehatan $\left(X\_{1}\right)$, angka partisipasi sekolah $\left(X\_{2}\right)$ indeks pembangunan manusia $\left(X\_{3}\right)$, pengangguran terbuka $\left(X\_{4}\right)$, sumber air minum $\left(X\_{5}\right)$ dan sanitasi layak $\left(X\_{6}\right)$. Eksperimen dilakukan dengan melakukan percobaan sebanyak 10 iterasi, dimana model terbaik diperolah pada iterasi ke-4 yakni dengan tingkat akurasi rata-rata 86%, dimana proses ini juga menemukan bahwa faktor yang berpengaruh terhadap kemiskinan yakni $X\_{3},X\_{4},X\_{5},$ dan $X\_{6}$. Pendekatan ini memiliki rata-rata akurasi yang baik, namun teknik ini perlu menentukan parameter jumlah iterasi yang diperlukan untuk mendapat hasil terbaik.

Ermawati [3] menggunakan klasifikasi penduduk miskin untuk menentukan kelayakan mendapatkan program bantuan pangan non tunai (BPNT) dengan menggunakan C4.5 dan C4.5 berbasis *particleswarn optimization* (PSO). Hasil eksperiment ini menunjukan tingkat akurasi 98.56 dan 98.92 untuk C4.5 dan C4.5 berbasis *particle swarm optimization*. Kumila dkk [5] menggunakan dua buah metode yakni *Moving Average* dan Naïve Bayes untuk kemiskinan Tenggara Barat (NTB) dengan data tahun 2002-2018 untuk memprediksi data tahun 2019 diperoleh bahwa metode Naïve Bayes memiliki hasil yang lebih baik dengan nilai MAD, MSE, RMSE dan MAPE berturut-turut yakni 41.427,188; 2.711.468.146; 52.071,760; dan 0.043. Syahrudin dkk [4] menggunakan *back propagation* untuk memprediksi angka kemiskinan di Indonesia. Pada penelitian ini menggunakan data dari tahun 2012-2019, dan memperoleh data hasil prediksi tahun 2020 bahwa penduduk miskin pada tahun tersebut ialah 332,005 jiwa. Adapun untuk pengukuran metode yang dilakukan menggunakan MSE dan MAPE dengan nilai masing-masing ialah 0.119 dan 2.298.

* 1. Metode *Rough Set*

*Rough set* merupakan *mathematic tool* yang sangat handal, ia dapat menganai data yang *vague, incomplete* dan *inconsistent*. *Rough set* telah digunakan dalam banyak bidang seperti Rezki [11] menggunakan *rough set* digunakan untuk menganalisa kinerja guru, Jamaris [12] menggunakan *rough set* untuk sistem pengambil keputusan kelayakan bantuan hibah fasilitas rumah ibadah, Sembiring dan Manurung [13] menggunakan *rough set* pada bidang ekonomi untuk analisis capaian keuntungan perusahaan, Sihombing [14] menggunakan rough set pada analisa bencana yakni analisis dampak tanah longsor.

Pada bidang pendidikan *rough set* digunakan untuk prediksi kelulusan mahasiswa pada uji kompetensi [15], sedangkan Prajana [16] menggunakan *rough set* untuk memprediksi kelulusan siswa pada ujian nasional. Kemudian, Putra dkk [17] menggunakan *rough set* untuk memprediksi kecerdasan anak. Raharjo dan Windarto [18] menggunakan *rough set* untuk memprediksi tingkat pemahaman mahasiswa terhadap mata kuliah.

Model *Rough Set* ini mempunyai beberapa kelebihan dibandingkan dengan yang lain yaitu: menyediakan algoritma yang efisien untuk menemukan pola yang tersembunyi dalam data, menemukan reduksi dari himpunan data, mengevaluasi signifikansi data, menghasilkan himpunan aturan-aturan keputusan dari data, mudah untuk dimengerti, menawarkan interpretasi yang mudah dari hasil, serta dapat digunakan untuk data kualitatif maupun kuantitatif [7].

* + 1. *Information System* dan *Decision Information System*

Pada rough set, data disajikan dalam tabel sistem informasi atau disebut dengan *information system*. Suatu *Information System* diformulasikan dengan $IS=(U, A, V, f)$, dimana $U$ ialah himpunan semesta yang berisi objek, dimana $U\ne ∅$, $A$ ialah himpunan atribut, dimana $A\ne ∅$. Sedangkan $V=\bigcup\_{a\in A}^{}V\_{a}$, dimana $V\_{a}$ ialah domain untuk atribut $a$, dan $f:U×A$ ialah fungsi yang memetakan antara nilai dari suatu objek $x\in U$ dan $a\in A$. Sedangkan *Information Decision Sistem* diformulasikan dengan $DIS=(U, A∪\{d\}, V,f)$, dimana $U$ ialah himpunan semesta yang berisi objek, dimana $U\ne ∅$, $A$ ialah himpunan atribut, dimana $A\ne ∅$, $d$ ialah atribut keputusan (*decision attribute*), $d\notin A$. Sedangkan $V=\bigcup\_{a\in A∪\{d\}}^{}V\_{a}$, dimana $V\_{a}$ ialah domain untuk atribut $a$, dan $f:U×A$ ialah fungsi yang memetakan antara nilai dari suatu objek $x\in U$ dan $a\in A∪\{d\}$.

* + 1. *Indiscernible Relation*

Misalkan $U$ ialah himpunan semesta, dimana $U\ne ∅$, sedangkan $A$ ialah himpunan atribut, dimana $A\ne ∅$, $S=\left(U,A\right)$ ialah *information system*. *Indiscernible Relation* ialah relasi biner antar dua buah objek $x, y\in U$ dimana untuk setiap himpunan atribut $B⊆A$ maka $∀b\in B\rightarrow b\left(x\right)=b\left(y\right)$. *Indescernible Relation* terhadap atribut $B$ umumnya diformulasikan sebagai berikut:

|  |  |
| --- | --- |
| $$IND\left(B\right)=\left\{∀b\in B, b\left(x\right)=b\left(y\right)\right\}$$ | (1) |

* + 1. *Set Approximation*

Misalkan $U$ ialah himpunan semesta, dimana $U\ne ∅$ dan himpunan bagian $X⊆U$, maka *set approximation* terdiri dari *lower approximation* dan *upper approximation* keduanya masing-masing diformulasikan oleh (2) dan (3).

|  |  |
| --- | --- |
| $$\overline{BX}=\bigcup\_{p\in IND\left(B\right)}^{}\left\{x|x\in p Λ p⊆X\right\}$$ | (2) |
| $$\overbar{BX}=\bigcup\_{p\in IND\left(B\right)}^{}\left\{x|x\in p Λ p∩X\ne ∅\right\}$$ | (3) |

Dari persamaan (2) dan persamaan (3) kita dapat menghitung *roughness accuracy* dengan menggunakan formula sebagai berikut:

|  |  |
| --- | --- |
| $$α\_{B}=\frac{\overline{X\left(B\right)}}{\overbar{X\left(B\right)}}$$ | (4) |

Nilai $α\_{B}$ berada pada rentang $0\leq α\_{B}\leq 1$. $α\_{B}$ yang menunjukan kesamaran (*vague*). Jika $α\_{B}=1$, maka ia bersifat *crisp (precise)/*tepat, dan samar jika $α\_{B}<1$.

*Lower Approximation* $\overline{X\left(B\right)}$ ialah area dimana suatu objek dapat dengan jelas apakah ia anggota dari $X$. *Lower approximation* juga disebut juga dengan *Positive Region*. Dari sini kita dapat menentukan *dependency attribute* dengan menggunakan (5).

|  |  |
| --- | --- |
| $$γ\left(C,D\right)= \frac{POS\_{B}}{\left|U\right| }=\bigcup\_{X⊆U}^{}\frac{\overline{X\left(B\right)}}{|U|}$$ | (5) |

* + 1. *Reduct*

Misalkan $U$ ialah himpunan semesta, dan $U\ne ∅$, $A$ ialah himpunan atribut, dimana $A\ne ∅$. Misalkan $B⊆A$ dan $b\in B$, maka atribut $b$ dikatakan $dispensible$(dapat dibuang) jika $γ(B, D)=γ(B-\left\{c\right\},D)$. Sebaliknya jika $γ(B, D\ne γ(B-\left\{c\right\},D)$maka $b$ *indispensible* (tidak dapat dibuang).

1. Metode Penelitian

Penelitian ini dilaksanakan di Provinsi Kalimantan Timur. Penelitian ini merupakan penelitian kuantitatif dengan menggunakan data sekunder. Sumber data penelitian ini adalah Badan Pusat Statistik (BPS) Provinsi Kalimantan Timur.Atribut/variabel yang digunakan dalam penelitian ini adalah data kemiskinan, Produk Domestik Regional Bruto (PDRB), Indeks Pengembangan Manusia (IPM), Inflasi dan pengangguran. Periode waktu data yang digunakan adalah data selama 15 tahun dari tahun 2006 – 2020.

Data yang telah diperoleh kemudian diolah dengan menggunakan metode rough set. Tahapan-tahapan pengolahan data pada metode ini adalah (1) Pembentukan tabel sistem informasi, (2) Mengelompokkan data yang memiliki nilai atribut yang sama, (3) Menentukan *lower approximation* dan *upper approximation*, (4) Menentukan *dependency attribute*, (5) Melakukan reduksi yakni menghapus atribut yang berlebihan, (6) menghasilkan rule/aturan dengan menggunakan atribute yang tersisa.

1. Hasil dan Pembahasan
	1. Koleksi Data

Lima *attribute*/ variabel yang digunakan dalam penelitian ini yakni (1) Pengangguran, (2) Produk Domestik Regional Bruto (PDRB), (3) Indeks Pengembangan Manusia (IPM), (4) Inflasi dan (5) Kemiskinan. Pada penelitian ini menggunakan data dari BPS provinsi Kalimantan Timur dari tahun 2006 hingga 2020 seperti yang ditunjukan pada Tabel 1.

**Tabel 1** : Data Ekonomi Kalimantan Timur

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Objek** | **Tahun** | **Pengangguran** | **PDRB** | **IPM** | **Inflasi** | **Kemiskinan** |
| P1 | 2006 | 13.43 | 2.88 | 73.26 | 6.04 | 11.41 |
| P2 | 2007 | 12.07 | 1.84 | 73.77 | 8.3 | 11.04 |
| P3 | 2008 | 11.11 | 4.9 | 74.52 | 13.06 | 9.51 |
| P4 | 2009 | 10.83 | 2.28 | 75.11 | 4.31 | 7.73 |
| P5 | 2010 | 10.1 | 5.1 | 71.31 | 7.28 | 7.66 |
| P6 | 2011 | 9.84 | 6.3 | 72.02 | 6.35 | 6.63 |
| P7 | 2012 | 8.9 | 5.26 | 72.62 | 5.6 | 6.38 |
| P8 | 2013 | 8.04 | 2.25 | 73.21 | 9.65 | 6.38 |
| P9 | 2014 | 7.38 | 1.71 | 73.82 | 7.66 | 6.31 |
| P10 | 2015 | 7.5 | -1.2 | 74.17 | 4.89 | 6.23 |
| P11 | 2016 | 7.95 | -0.38 | 74.59 | 3.39 | 6.11 |
| P12 | 2017 | 6.91 | 3.13 | 75.12 | 3.15 | 6.19 |
| P13 | 2018 | 6.6 | 2.64 | 75.83 | 3.24 | 6.03 |
| P14 | 2019 | 6.09 | 4.74 | 76.61 | 1.66 | 5.94 |
| P15 | 2020 | 6.87 | 2.85 | 76.24 | 0.25 | 6.1 |

Sumber: Badan Pusat Statistik Provinsi Kalimatan Timur

* 1. *Pre-processing*

Pada tahap ini, data numerik pada Tabel 1, ditransformasi menjadi data diskrit dengan membagi menjadi beberapa kuartil dengan menggunakan persamaan berikut:

|  |  |
| --- | --- |
| $$Q\_{i}=\frac{i(n+1)}{4}$$ | (6) |

Dimana $Q\_{i}$ ialah kuartil ke-$i$ dengan $i=\{1,2,…,m\}$, $m$ ialah jumlah kuartil dan $n$ ialah jumlah objek pengamatan. Pada penelitian ini, data dibagi menjadi tiga kuartil yang masing-masing sebagai berikut:

1. $Q\_{1}$ = Tinggi
2. $Q\_{2}$ = Sedang
3. $Q\_{3}$ = Rendah
4. $Q\_{4}$ = Sangat Rendah
	1. Model *Rough Set*
		1. Sistem Informasi Keputusan

Berdasarkan langkah pada sesi 4.2, sistem informasi keputusan dari data ialah ditunjukan pada Tabel 2. Tabel 2 merupakan hasil dari transformasi data, yakni dari data kontinue menjadi data diskrit dengan aturan pengambilan keputusannya menggunakan formula (6) untuk membagi data menjadi empat kriteria golongan.

**Tabel 2 :** Transformasi Data Tabel 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Objek** | **Pengangguran** | **PDRB** | **IPM** | **Inflasi** | **Kemiskinan** |
| P1 | tinggi | sedang | rendah | Sedang | tinggi |
| P2 | tinggi | sangat rendah | rendah | Tinggi | tinggi |
| P3 | tinggi | sedang | sedang | Tinggi | tinggi |
| P4 | sedang | rendah | sedang | Rendah | sedang |
| P5 | sedang | tinggi | sangat rendah | Sedang | sedang |
| P6 | sedang | tinggi | sangat rendah | Sedang | sedang |
| P7 | sedang | tinggi | sangat rendah | Rendah | sedang |
| P8 | rendah | rendah | sangat rendah | tinggi | rendah |
| P9 | rendah | sangat rendah | rendah | sedang | rendah |
| P10 | rendah | sangat rendah | rendah | rendah | rendah |
| P11 | rendah | sangat rendah | sedang | rendah | rendah |
| P12 | sangat rendah | sedang | sedang | sangat rendah | sangat rendah |
| P13 | sangat rendah | rendah | tinggi | sangat rendah | sangat rendah |
| P14 | sangat rendah | sedang | tinggi | sangat rendah | sangat rendah |
| P15 | sangat rendah | rendah | tinggi | sangat rendah | sangat rendah |

* + 1. *Indiscernible Relation*

Dengan menggunakan (1), didapat relasi ekuivalensi untuk atribut/variabel independen $B=\{Pengangguran, PDRB, IPM, Inflasi\}$, maka diperoleh

$$IND\left(B\right)=\left\{\left\{P1\right\},\left\{P2\right\},\left\{P3\right\},\left\{P4\right\},\left\{P5,P6\right\},\left\{P7\right\},\left\{P8\right\},\left\{P9\right\},\left\{P10\right\},\left\{P10\right\},\left\{P11\right\},\left\{P12\right\}, \left\{P13, P15\right\},\left\{P14\right\}\right\}$$

* + 1. *Set Approximation*

Setelah memperoleh Indiscerible relation dengan menggunakan relasi biner, maka dapat ditentukan set approximation dengan menggunakan (2), sehingga selanjutnya dapat dihitung r*oughness accuracy* dan *dependency attribute*. Adapun *lower approximation* and *upper approximation* untuk masing-masing kelas ialah sebagai berikut:

1. Kelas Tinggi. *Lower Approximation* $\overline{X\left(B\right)}=\left\{P1, P2,P3\right\}$ dan $\overline{X(B)}=\left\{P1, P2,P3\right\}$
2. Kelas Sedang. *Lower Approximation* $\overline{X\left(B\right)}=\left\{P4, P5,P6, P7\right\}$ dan $\overline{X(B)}=\left\{P4, P5,P6, P7\right\}$
3. Kelas Rendah. *Lower Approximation* $\overline{X\left(B\right)}=\left\{P8,P9,P10,P11\right\}$ dan $\overline{X(B)}=\left\{P8,P9,P10,P11\right\}$
4. Kelas Sangat Rendah. *Lower Approximation* $\overline{X\left(B\right)}=\left\{P12,P13,P14,P15\right\}$ dan $\overline{X(B)}=\left\{P12,P13,P14,P15\right\}$

Dari *Lower approximation* and *upper approximation* untuk setiap kelas maka dapat diketahui nilai *roughness accuracy* untuk masing-masing kelas. Dari (4) *roughness accuracy* ($α\_{B}$)untuk Tinggi, Sedang, Rendah dan sangat rendah masing-masing ialah 1, 1, 1 dan 1. Dari sini kita simpulkan bahwa kelas bersifat *crisp* terhadap atribut $B$. Kemudian dengan (5) dapat dihitung *depency attribute* ($γ\_{B}$) $B$ terhadap keputusan yakni 1, artinya tabel sistem informasi atau data bersifat konsisten, sehingga kita dengan mengetahui nilai ini kita dapat simpulkan bahwa $B$ dapat dengan tepat (*crisp*/*precise*) mendefinisikan setiap kelas pada tabel sistem informasi.

* + 1. *Reduct*

Reduct yakni melakukan reduksi terhadap parameter yang telah digunakan sebelumnya, dengan membuah parameter yang *superflous* (berlebihan). Pada eksperimen ini, reduksi parameter menggunakan metode yang diusulkan oleh Pawlak. Disini kita menggunakan *brute force* yakni dengan mencoba semua kemungkinan subset atribut $B⊆A$ yang memiliki $γ\_{B}=1$, sehingga didapatkan nilai seperti ditunjukan pada Tabel 3.

**Tabel 3 :**  Nilai *Dependency Attribute* untuk masing-masing subset atribut$ B⊆A$

|  |  |
| --- | --- |
| **Sub set atribut** | ***Depency Attribute* (**$γ\_{B}$**)** |
| Inflasi | 0.266667 |
| IPM | 0.200000 |
| Pengangguran | 1.000000 |
| IPM, Inflasi | 0.733333 |
| PDRB | 0.200000 |
| PDRB, Inflasi | 1.000000 |
| PDRB, IPM | 0.666667 |
| PDRB, IPM, Inflasi | 1.000000 |
| Pengangguran, Inflasi | 1.000000 |
| Pengangguran, IPM | 1.000000 |
| Pengangguran, IPM, Inflasi | 1.000000 |
| Pengangguran, PDRB | 1.000000 |
| Pengangguran, PDRB, Inflasi | 1.000000 |
| Pengangguran, PDRB, IPM | 1.000000 |
| Pengangguran, PDRB, IPM, Inflasi | 1.000000 |

Dari Tabel 3, didapatkan $γ\_{\left\{pengangguran\right\}}=γ\_{B}$, dimana $B=Pengangguran, PDRB, IPM, Inflasi$ yang bernilai 1. Yang mana artinya bila kita hanya menggunakan atribute pengangguran dalam menganalisa kemiskinan maka hasilnya akan sama halnya dengan menggunakan atribute $B$ (Pengangguran, PDRB, IPM dan Inflasi).

* + 1. *Generate Rule*

Dari proses *reduct* diperoleh bahwa atribut pengangguran adalah atribut yang *indispensible* dan memiliki nilai *attribute dependency* yang sama dengan $B$, dari sini kita dapat merepresentasikan ulang Tabel 2 hanya dengan menggunakan atribut pengangguran seperti ditunjukan pada Tabel 4.

**Tabel 4** : Kelas Ekuivalensi

|  |  |  |
| --- | --- | --- |
| **Objek** | **Pengangguran** | **Kemiskinan** |
| EC1 | tinggi | tinggi |
| EC2 | sedang | sedang |
| EC3 | rendah | rendah |
| EC4 | sangat rendah | sangat rendah |

Dari Tabel 4 dapat dibuat sebuah hubungan antara independen atribut terhadap dependen atribut dalam bentuk aturan *if..then* sebagai berikut:

1. If Pengangguran=Tinggi then Kemiskinan Tinggi
2. If Pengangguran=Sedang then Kemiskinan Sedang
3. If Pengangguran=Rendah then Kemiskinan Rendah
4. If Pengangguran=Sangat Rendah then Kemiskinan Sangat Rendah

Dari hasil pembahasan diatas ditemukan bahwa atribut pengangguran memiliki *dependency attribute* yakni 1 atau jika dikonversi menjadi persen menjadi 100%. Hal ini memiliki makna bahwa pengangguran merupakan faktor utama yang sangat berpengaruh terhadap kemiskinan di Provinsi Kalimantan Timur. Sehingga jika ingin melihat prediksi atau perkembangan kemiskinan ke depan maka dapat melihat dari angka penganggurannya. Jika dilihat dalam bentuk grafik seperti pada gambar 1.



**Gambar 1:** Visualisasi Data

 Gambar 1 menampilkan hasil visualisasi data pengangguran, PDRB, IPM, inflasi dan kemiskinan. Dari gambar dapat dilihat bahwa garis untuk pengangguran hampir sama bentuknya dengan garis kemiskinan.

1. Kesimpulan

Berdasarkan hasil analisis dan perhitungan dengan menggunakan metode rough set untuk melakukan analisa kemiskinan di Provinsi Kalimantan Timur, maka dapat disimpulkan bahwa pengangguran adalah faktor utama yang mempengaruhi terjadinya kemiskinan di Provinsi Kalimantan Timur, dimana faktor ini memberikan nilai *dependency attribute* sebesar 100%, sehingga dengan kata lain untuk menurunkan tingkat kemiskinan pada provinsi Kalimantan Timur, maka perlu adanya upaya untuk menyediakan lapangan kerja untuk menurunkan pengangguran.

Daftar Pustaka

[1] A. Wanto, “Penerapan Jaringan Saraf Tiruan Dalam Memprediksi Jumlah Kemiskinan Pada Kabupaten/Kota Di Provinsi Riau,” *Klik - Kumpul. J. Ilmu Komput.*, vol. 5, no. 1, p. 61, 2018, doi: 10.20527/klik.v5i1.129.

[2] S. Mamase and R. S. Sinukun, “Prediksi Tingkat Kemiskinan Provinsi Gorontalo Menggunakan Metode Gabungan K-Means dan Generalized Regression Neural Network,” *J. ENERGY*, vol. 10, no. 2, pp. 29–34, 2018.

[3] E. Ermawati, “Algoritma Klasifikasi C4.5 Berbasis Particle Swarm Optimization Untuk Prediksi Penerima Bantuan Pangan Non Tunai,” *Sistemasi*, vol. 8, no. 3, p. 513, 2019, doi: 10.32520/stmsi.v8i3.576.

[4] Syaharuddin, E. Pujiana, I. P. Sari, V. M. Mardika, and M. Putri, “Analisis Algoritma Back Propagation Dalam Prediksi Angka Kemiskinan Di Indonesia,” *J. Pendidik. Berkarakter*, vol. 3, no. 1, pp. 11–17, 2020, [Online]. Available: http://journal.ummat.ac.id/index.php/pendekar/article/view/2814.

[5] A. Kumila, B. Sholihah, E. Evizia, N. Safitri, and S. Fitri, “Perbandingan Metode Moving Average dan Metode Naïve Dalam Peramalan Data Kemiskinan,” *JTAM | J. Teor. dan Apl. Mat.*, vol. 3, no. 1, p. 65, 2019, doi: 10.31764/jtam.v3i1.764.

[6] R. Efendi, V. A. Dewi, S. Basriati, and D. S. Ss, “Pengaruh Pengangguran dan PDRB Terhadap Tingkat Kemiskinan Menggunakan Regresi Linier Berganda dan Rough Sets,” no. November, pp. 651–657, 2018.

[7] Z. Pawlak, “Rough sets,” *Int. J. Comput. Inf. Sci.*, vol. 11, no. 5, pp. 341–356, 1982, doi: 10.1007/BF01001956.

[8] M. Kryszkiewicz, “Rough set approach to incomplete information systems,” *Inf. Sci. (Ny).*, vol. 112, no. 1–4, pp. 39–49, 1998, doi: 10.1016/S0020-0255(98)10019-1.

[9] R. S. S. Saprina Mamase, “Prediksi Tingkat Kemiskinan Provinsi Gorontalo dengan Metode GRNN,” in *Seminar Nasional Humaniora & Aplikasi Teknologi Informasi 2018 (SEHATI 2018) PREDISKI*, 2018, pp. 29–32.

[10] V. Rahmadi, R. Yulistiani, R. Sheffi, T. Gultom, and M. M. Santoni, “Prediksi Penyebab Utama Kemiskinan Di Indonesia Dengan,” in *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, 2020, pp. 172–181.

[11] D. A. Rezki, “Analisa Kinerja Guru Sekolah di Dinas Pendidikan Dengan Menggunakan Metode Rought Set ( Studi Kasus : Dinas Pendidikan Provinsi Sumatera Utara ),” *J. Comput. Syst. Informatics*, vol. 1, no. 3, pp. 172–177, 2020.

[12] M. Jamaris, “Implementasi Metode Rough Set Untuk Menentukan Kelayakan Bantuan Dana Hibah Fasilitas Rumah Ibadah,” *INOVTEK Polbeng - Seri Inform.*, vol. 2, no. 2, p. 161, 2017, doi: 10.35314/isi.v2i2.203.

[13] M. A. Sembiring and Z. Azhar, “Implementasi Metode Rough Set Untuk Menganalisa Laba/Rugi Pada Suatu Perusahaan Distributor (Studi Kasus : Usaha Kita Ps Payakumbuh),” pp. 1–6, 2013.

[14] P. Sihombing, “Implementasi Metode Rough Set Dalam Memprediksi Dampak Tanah Longsor ( Studi Kasus Badan Penanggulangan Bencana Daerah ( BPBD ) Provinsi Sumatera Utara a : U,” vol. 6, no. 4, pp. 407–415, 2019.

[15] K. Suryani, “Prediksi Peluang Kelulusan Mahasiswa PTIK dalam Uji Kompetensi Microsoft Office 2010 menggunakan Teori Rough Set,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 1–10, 2016, doi: 10.25077/teknosi.v2i1.2016.1-10.

[16] A. Prajana, F. Sains, T. Universitas, I. Negeri, A. Raniry, and B. Aceh, “Penerapan Teory Rough Set Untuk Memprediksi Tingkat Kelulusan Siswa Dalam Ujian Nasional Pada Sma Negeri 5 Kota Banda Aceh,” *J. Islam. Sci. Technol.*, vol. 2, no. 1, pp. 75–88, 2016, [Online]. Available: www.jurnal.ar-raniry.com/index.php/elkawnie.

[17] A. Putra, Z. A. Matondang, N. Sitompul, I. Pendahuluan, and A. Prediksi, “Implementasi Algoritma Rough Set Dalam Memprediksi Kecerdasan Anak,” *J. Pelita Inform.*, vol. 7, no. 2, pp. 149–156, 2018.

[18] M. R. Raharjo and A. P. Windarto, “Penerapan Machine Learning dengan Konsep Data Mining Rough Set ( Prediksi Tingkat Pemahaman Mahasiswa terhadap Matakuliah ),” vol. 5, pp. 317–326, 2021, doi: 10.30865/mib.v5i1.2745.