

Artikel : [Akses terbuka/Open Access](#)

PERBANDINGAN ALGORITMA MACHINE LEARNING UNTUK PENENTUAN KLASIFIKASI KEMISKINAN MULTIDIMENSI DI PROVINSI NUSA TENGGARA TIMUR

Sitasi : Utomo. 2022, JSTAR 2(1), 36-46.

Kronologi naskah.

Submit : 3 Maret 2022

Revisi : 22 April 2022

Diterima : 22 April 2022



Penyedia Data Statistik Berkualitas untuk
Indonesia Maju

REFORMASI BIROKRASI



PERBANDINGAN ALGORITMA *MACHINE LEARNING* UNTUK PENENTUAN KLASIFIKASI KEMISKINAN MULTIDIMENSI DI PROVINSI NUSA TENGGARA TIMUR

Kristanto Setyo utomo

Fungsi Neraca Wilayah dan Analisis Statistik, Badan Pusat Statistik Provinsi NTT, Indonesia
*korespondensi *author*: kristanto_su@bps.go.id.

Abstract

The Covid-19 pandemic has proven to directly impact the percentage of poverty in the Province of East Nusa Tenggara. However, the determination of the size of poverty so far has mostly been carried out using an economic dimension approach, namely the poverty line. This study classifies multidimensional poverty, namely the dimensions of health, education, economy, and standard of living. In this multidimensional poverty classification, this research utilizes machine learning algorithms. The test results show that the Decision Tree algorithm is the best algorithm for classifying multidimensional poverty in East Nusa Tenggara Province with an accuracy rate of 82.69 percent, precision of 84.08 percent, and recall of 97.56 percent. This algorithm, furthermore, shows that the birth attendant indicator in the health dimension and primary education in the education dimension have high gain values. Therefore, these two indicators become the primary decision node in the Decision Tree, which signifies that they share the highest contribution to multidimensional poverty and, as such, should be a high policy priority of the government of Nusa Tenggara Timur Province.

Keyword: *Algoritma Klasifikasi, Kemiskinan Multidimensi, Machine Learning*

1. Pendahuluan

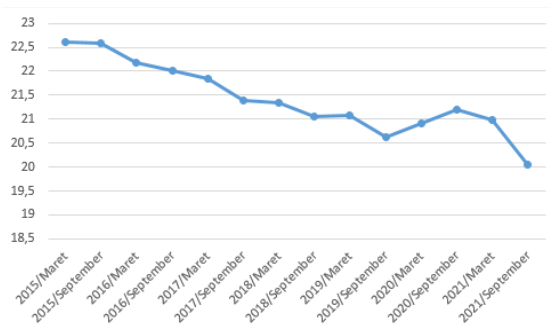
Isu kemiskinan telah lama menjadi persoalan mendasar yang menjadi pusat perhatian pemerintah. Pemerintah telah banyak melakukan berbagai kebijakan dalam menanggulangi permasalahan ini (Alfista dkk, 2019). Namun penyebaran pandemi Covid-19 yang mulai terjadi pada awal tahun 2020 menjadi tantangan baru bagi pemerintah dalam menyelesaikan masalah kemiskinan ini.

Upaya pemerintah dalam menghambat penyebaran pandemi

Covid-19 berdampak pada peningkatan jumlah orang miskin baru dan tentunya berimplikasi pada peningkatan angka kemiskinan secara total (Suharyadi dkk, 2020 dalam Tarigan dkk, 2020). Berdasarkan penelitian Setyadi & Indriyani (2021), menunjukkan bahwa peningkatan jumlah kasus positif Covid-19 di Indonesia terbukti berpengaruh positif dan signifikan terhadap kemiskinan. Selanjutnya, hasil penelitian Tarigan dkk, (2020) menunjukkan bahwa jumlah terbesar perubahan status penduduk dari kategori rentan

miskin menjadi miskin terdapat di Pulau Jawa, di mana kasus terkonfirmasi positif Covid-19 cukup tinggi.

Provinsi Nusa Tenggara Timur juga tidak terhidar dari dampak buruk pandemi Covid-19 terhadap kondisi kemiskinan ini. Pada Grafik 1 dapat dilihat bahwa dari tahun 2015 persentase kemiskinan di provinsi ini sudah mengalami *trend* penurunan, namun akibat pandemi Covid-19 persentase angka kemiskinan Maret 2020 kembali mengalami peningkatan sebagai akibat diberlakukannya berbagai kebijakan pembatasan oleh pemerintah. Namun pada September 2020 laju kemiskinan di Indonesia kembali mengalami penurunan sejalan dengan mulai dilonggarkannya kebijakan pembatasan mobilitas penduduk serta adanya berbagai intervensi kebijakan pemerintah dalam mengatasi dampak pandemi Covid-19 di Indonesia (BPS, 2022).



Grafik 1. Persentase Kemiskinan di Provinsi Nusa Tenggara Timur tahun 2015-2021.

Tersedianya data kemiskinan yang cepat dan akurat menjadi salah satu aspek penting dalam mendukung strategi penanggulangan kemiskinan. Dengan adanya data tersebut maka

pemerintah juga akan dapat menentukan kebijakan yang harus dilakukan untuk penanggulangan kemiskinan dengan cepat dan tepat (Ferezagia, 2018). Untuk menjawab tantangan ini, terdapat berbagai literatur metode yang memanfaatkan kecerdasan buatan (*artificial intelligence*) untuk memprediksi kemiskinan dengan cepat salah satunya menggunakan *machine learning* (Sani dkk, 2018). *Machine learning* (ML) sudah digunakan secara luas di berbagai bidang kehidupan seperti bidang ekonomi, telekomunikasi, kesehatan dan pengobatan, pendidikan dan kemiskinan (Kaunang, 2019). Implementasi ML ini dilakukan dengan melihat pola kumpulan data tertentu tanpa instruksi berdasarkan aturan pemrograman yang eksplisit (Sani dkk, 2018).

Penelitian ini membandingkan pengujian tiga algoritma ML untuk menentukan klasifikasi kemiskinan multidimensional di Provinsi Nusa Tenggara Timur. Dari hasil pengujian ini akan didapatkan algoritma dari ML terbaik yang dapat diimplemestasikan untuk penentuan klasifikasi kemiskinan multidimensi yang terjadi di wilayah ini. Dengan demikian, pemerintah dapat menentukan kebijakan yang lebih terarah dan tepat sasaran untuk mengurangi kemiskinan di Provinsi Nusa Tenggara Timur.

2. Metodologi

Metode penelitian pada studi ini mengadopsi metode penelitian yang dilakukan oleh Yunita (2017), yakni menggunakan model “*Cross Standard Industry for Data Mining (CRISP-DM)* yang terdiri dari 6 tahap, yaitu: *business understanding, data understanding, data preparation, modeling, evaluation, deployment*”.

Tahap *Business Understanding, Data Understanding* dan *Preparation*

Pengukuran kemiskinan makro di Indonesia sampai saat ini masih menggunakan pendekatan moneter (Sumargo dkk, 2019). Pendekatan dimensi ekonomi dalam mendefinisikan kemiskinan ini lebih sering digunakan secara umum. Kemudahan variabel pada dimensi ekonomi untuk dilakukan pengamatan dan pengukuran serta kemudahan dalam melihat keterbandingan antara penduduk menjadi faktor pendukung seringnya dimensi ekonomi ini digunakan dalam pengukuran kemiskinan (Fadlillah dkk, 2016). Namun Budiantoro dkk (2013), menyimpulkan bahwa menelaah kemiskinan secara multidimensional sangat diperlukan untuk merumuskan kebijakan pengentasan kemiskinan agar lebih tepat sasaran. Oleh karena itu, UN *Development Programme (UNDP)* dan *Oxford Poverty and Human Development Initiative (OPHI)* telah memformulasikan dimensi kemiskinan menurut metode Alkire-Foster (Alkire, 2013). Metode Alkire-Foster tersebut juga pernah diterapkan di Indonesia

untuk menghasilkan dimensi dan indikator yang menggambarkan kondisi kemiskinan multidimensi (Sumargo & Simajuntak 2019). Selain itu, Fadilah dkk, (2016) juga pernah melakukan pengujian keterkaitan kemiskinan multi dimensi dengan dimensi Indeks Pembangunan Manusia (IPM). IPM merupakan indeks yang mampu menggambarkan kualitas hidup manusia sehingga sangat erat kaitannya dengan kemiskinan multidimensi.

Berdasarkan studi literatur tersebut, maka penentuan dimensi dan indikator untuk menentukan multidimensi kemiskinan pada penelitian ini dapat dilihat pada Tabel 1. Sedangkan penentuan indikator pada masing-masing dimensi dilakukan dengan studi literatur yang menunjukkan adanya korelasi antara setiap dimensi dengan indikator didalamnya. Salah satunya adalah pada dimensi kesehatan, pada dimensi ini terdapat penambahan indikator kepemilikan jaminan kesehatan. Penelitian Djunawan (2019), menunjukkan bahwa kepemilikan jaminan kesehatan bagi masyarakat miskin khususnya di daerah perkotaan berpengaruh terhadap pemanfaatan fasilitas kesehatan yang disediakan pemerintah. Oleh karena itu, pemberian subsidi pemerintah kepada masyarakat miskin melalui jaminan kesehatan ini diharapkan mengurangi kemiskinan dari dimensi kesehatan.

Indikator jenis pekerjaan kepala rumah tangga dan kepemilikan tabungan pada penelitian ini digunakan

untuk memberikan gambaran kemiskinan dari dimensi ekonomi. Berdasarkan penelitian Sari & Purwanti (2012) menunjukkan bahwa jenis pekerjaan kepala rumah tangga berpengaruh signifikan terhadap terjadinya kemiskinan. Begitu juga dengan kepemilikan tabungan yang terbukti memiliki pengaruh signifikan terhadap nilai kerentanan kemiskinan pada suatu rumah tangga (Adnyani & Sugiharti, 2019).

Tabel 1. Indikator Menurut Dimensi

Dimensi	IPM	MPI UNDP dan OPHI	MPI Indonesia	Studi ini
Kesehatan	Angka Harapan Hidup	Gizi; Kematian Bayi	Imunisasi; Penolong Kelahiran	Kepemilikan Jaminan Kesehatan; Imunisasi Bayi; Penolong Kelahiran
Pendidikan	Angka Harapan Lama Sekolah	Lama Sekolah; Kehadiran Dalam Sekolah	Lama Sekolah; Partisipasi Sekolah; Melek Huruf	Pemenuhan Pendidikan Dasar; Melek Huruf
Ekonomi	Pengeluaran perkapita yang disesuaikan	-	-	Pekerjaan KRT; Kepemilikan Rekening tabungan
Standar Hidup Layak	-	Bahan Bakar Memasak; Sanitasi; Air Bersih; Listrik; Perumahan; Kepemilikan Aset	Jenis Lantai; Jenis Dinding; Jenis Atap; BBM; Listrik; Sumber Air Minum; Sanitasi	Kepemilikan Aset; Keadaan Rumah Layak Huni

Sumber: Alkire dkk, (2013), Budiantoro dkk, (2013), Fadilah dkk, (2016), Sumargo & Simajuntak (2019).

Sumber data yang digunakan pada penelitian ini berupa data primer yang diperoleh dari pendataan SUSENAS tahun 2021 yang dilakukan

Badan Pusat Statistik. SUSENAS merupakan survei yang dirancang untuk mengumpulkan data sosial kependudukan yang relatif sangat luas. Data yang dikumpulkan antara lain menyangkut bidang pendidikan, kesehatan/gizi, perumahan, sosial ekonomi lainnya.

Tahap *Modeling, Evaluation, dan Deployment*

Kajian dalam penelitian ini menggunakan algoritma machine learning untuk menentukan klasifikasi kemiskinan multidimensi. Menurut Jordan dkk, (2015), ML merupakan metode penelitian yang mengimplementasikan berbagai disiplin ilmu seperti matematika, statistika, basis data, pemrograman paralel secara simultan. Pendekatan ini ditandai dengan penggunaan algoritma untuk mengekstrak pengetahuan dari kumpulan data yang besar dan heterogen. ML melakukan pembelajaran langsung dari data dan memperkirakan fungsi matematika yang merepresentasikan dari beberapa input, atau belajar menghubungkan satu atau lebih input ke satu atau lebih output untuk dapat merumuskan prediksi pada data baru.

Pada penelitian ini dilakukan perbandingan tiga algoritma *machine learning* yang termasuk dalam kategori prediktif pada fungsi klasifikasi. Sartika & Sensuse,(2017) menjelaskan bahwa kategori prediktif digunakan untuk menyimpulkan pola dari prediksi data latih (*training*) sebagai pembelajaran,

yang selanjutnya pola tersebut digunakan pada data yang belum terprediksi (*data test*). Han dkk (2011) mendefinisikan fungsi klasifikasi sebagai bentuk analisis data untuk mengekstrak model yang akan digunakan untuk memprediksi label kelas tertentu. Kelas dalam klasifikasi ini merupakan atribut dalam satu set data yang merupakan variabel *independent* dalam statistik. Proses klasifikasi data dilakukan dalam dua tahap, yaitu tahap pembelajaran dan tahap pengklasifikasian. Tahap pembelajaran merupakan tahapan dalam pembentukan model klasifikasi, sedangkan tahap pengklasifikasian merupakan tahapan penggunaan model klasifikasi untuk memprediksi label kelas dari suatu data. Terdapat beberapa pilihan algoritma ML yang dapat digunakan dalam pengklasifikasian data, namun dalam penelitian ini hanya membandingkan tiga algoritma saja, yakni *Naïve Bayes*, *K-Nearest Neighbour*, dan *Decision Tree*.

Naïve Bayes

Naïve Bayes merupakan salah satu algoritma *supervised learning* dalam ML yang paling populer, algoritma ini menggunakan konsep teorema *Bayes* atau probabilitas untuk memprediksi kelas data. Olsen Delen, (2008) dalam Watratan & Moeis, (2020) menyebutkan bahwa, “Ciri utama dari *Naïve Bayes Classifier* ini adalah asumsi yang sangat kuat (naif) akan independensi dari masing-masing kondisi/kejadian”. Adapun rumus algoritma *Naïve Bayes* tersebut adalah sebagai berikut:

$$p^{(H|E)} = \frac{p^{(E|H)} \times p^{(H)}}{p^{(E)}} \quad (1)$$

Dengan penjelasan $p^{(H|E)}$ adalah probabilitas hipotesis H dapat terjadi jika *evidence* E terjadi, $p^{(E|H)}$ merupakan probabilitas *evidence* E jika hipotesis H terjadi, $p^{(H)}$ adalah probabilitas hipotesis H tanpa melihat *evidence* E, $p^{(E)}$ adalah probabilitas *evidence* E tanpa melihat apapun.

K-Nearest Neighbour

K-Nearest Neighbour adalah algoritma pengklasifikasian yang didasarkan pada analogi, yaitu membandingkan data uji dengan data latih yang berada dekat dan memiliki kemiripan dengan data uji tersebut. *K-Nearest Neighbour* akan mencari jarak yang paling dekat antara data uji dengan *K-Nearest Neighbour* pada data latih. Sehingga kemiripan data uji dengan data latih didasarkan pada jaraknya (Ayudhitama & Pujiyanto, 2020).

Tahapan pada algoritma ini diawali dengan menghitung nilai bobot dari setiap atribut pada kelas yang ada dan menghitung bobot dengan keanggotaan kelas yang ada pada data latih dengan rumus sebagai berikut:

$$w = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (2)$$

di mana w adalah rata-rata bobot setiap atribut, x_n adalah data atribut dan n menunjukkan jumlah data. Setelah didapatkan masing-masing bobot maka dilanjutkan dengan menghitung jarak *euclidian* dengan rumus:

$$d_i = \sum_{i=1}^p w(x_{2i} - x_{1i})^2 \quad (3)$$

di mana d_i menunjukkan nilai jarak, p merupakan dimensi data, x_{1i} adalah nilai data latih variable ke 1, dan x_{2i} adalah nilai data uji variable ke i.

Decision Tree

Menurut Sartika & Sensuse (2017), algoritma *decision tree* merupakan algoritma yang umum digunakan untuk melakukan klasifikasi pada ML. *Decision tree* mencari solusi klasifikasi dengan menjadikan kriteria sebagai *node* yang saling berhubungan membentuk seperti struktur pohon. *Decision tree* adalah model prediksi terhadap suatu keputusan menggunakan struktur hirarki atau pohon. Setiap pohon memiliki cabang, di mana cabang tersebut mewakili suatu atribut yang harus dipenuhi untuk menuju cabang selanjutnya hingga berakhir di daun. Tahapan algoritma ini dimulai dengan menghitung *entropy* untuk menentukan cabang pohon dengan rumus sebagai berikut:

$$Entropy(S) = \sum_{j=1}^k -p_j \log_2 p_j \quad (4)$$

di mana S menunjukkan himpunan kasus, k adalah jumlah partisi S , dan p_j merupakan probabilitas yang didapat dari hasil jumlah *true* atau *false* dibagi dengan total kasus. Tahapan berikutnya adalah menghitung nilai *gain*, dengan ketentuan bahwa nilai *gain* yang paling tinggi akan menjadi cabang pohon yang pertama dengan rumus sebagai berikut:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|s_i|}{|s|} \times Entropy(s_i) \quad (5)$$

di mana A menunjukkan atribut, n adalah jumlah partisi A , $|s_i|$ menunjukkan jumlah kasus pada partisi ke i , $|s|$ menunjukkan jumlah kasus dalam S .

Ketiga algoritma pada ML tersebut akan dibandingkan akurasinya menggunakan metode *confusion matrix*. Metode ini berupa Tabel yang didalamnya terdapat *record* perbandingan hasil klasifikasi data uji berdasarkan data latih dengan data sebenarnya.

Tabel 2. *Confusion Matrix*

		Nilai sebenarnya	
		TRUE	FALSE
Nilai Prediksi	TRUE	True Positif	False Positif
	FALSE	False Negatif	True Negatif

Pengujian *confusion matrix* tersebut pada umumnya dilakukan tiga metode pengujian, yaitu:

1. Accuracy

Accuracy merupakan metode pengujian algoritma ML berdasarkan tingkat kedekatan antara nilai prediksi dengan nilai sebenarnya. Rumus menghitung akurasi yaitu:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

2. Precision

Precision adalah metode pengujian algoritma ML dengan melakukan suatu perbandingan data benar yang diperoleh sistem dengan jumlah seluruh data yang diambil oleh sistem yang benar maupun salah. Rumus menghitung presisi yaitu:

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\%$$

3. Recall

Recall adalah metode pengujian algoritma ML dengan melakukan perbandingan jumlah data benar yang diperoleh sistem dengan jumlah seluruh data benar yang diambil atau tidak diambil oleh sistem. Rumus menghitung *recall* yaitu:

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\%$$

Semakin tinggi nilai hasil uji pada ketiga metode tersebut maka model yang dihasilkan semakin baik

3. Hasil dan Pembahasan

Business Understanding, Data Understanding dan Preparation

Pada tahap ini diawali dengan penentuan label kemiskinan pada data SUSENAS yang dalam penelitian ini menggunakan konsep garis kemiskinan (GK) untuk makanan dan bukan makanan pada Provinsi Nusa Tenggara Timur tahun 2021 yaitu sebesar Rp 437.606,-/kapita/bulan (BPS, 2022). Tahap selanjutnya dilakukan *data validation*, *data integration*, dan *tranformation* agar dapat dilakukan pengujian. Tahap data *validation* digunakan untuk mengidentifikasi dan menghapus data yang *outlier (noise)*, data yang tidak lengkap (*missing value*) dan atau data tidak konsisten. Sedangkan data *integration* dan *tranformation* digunakan untuk

meningkatkan akurasi dan efesiensi algoritma (Yunita, 2017). Oleh karena itu, data yang digunakan penelitian ini menggunakan *categorical* untuk model algoritma *Naïve Bayes* dan *Decision Tree*. Namun pada model algoritma *K-Nearest Neighbour* data tersebut ditranformasi menjadi *numerical*. Tahap berikutnya data label dan indikator masing-masing dimensi tersebut dilakukan pemisahan menjadi data latih (*training*) dan data tes (*testing*) secara random dan dilakukan pengujian *standard cross validation* untuk memprediksi *error rate* menggunakan *confusion matrix*. Hasil *confusion matrix* pada masing-masing model algoritma ML ditunjukkan pada Tabel 3.

Tabel 3. Hasil Penghitungan *Confusion Matrix* Algoritma ML

Algoritma	Confusion Matrix	true Tidak Miskin	true Miskin
Naïve Bayes	pred. Tidak Miskin	945.687.318	127.119.222
	pred. Miskin	157.071.290	102.644.989
Decision Tree	pred. Tidak Miskin	1.075.804.025	203.639.131
	pred. Miskin	26.954.583	26.125.081
K-Nearest Neighbour	pred. Tidak Miskin	173.580.378	28.641.958
	pred. Miskin	44.490.723	19.389.239

Dari hasil matrix tersebut maka dapat dilakukan pengujian klasifikasi untuk menganalisis seberapa baik *classifier* untuk dapat mengenali kemiskinan dari empat dimensi yang telah ditentukan sebelumnya. Hasil pengujian masing-masing model algoritma *classifier* tersebut ditunjukkan pada Tabel 4.

Tabel 4. Hasil Pengujian *Accuracy*,
Precision, *Recall* Algoritma ML

Algoritma	Accuracy	Precision	Recall
Naïve Bayes	78.67% +/- 1.31% (micro average: 78.67%)	88.15%	85.76%
Decision Tree	82.69% +/- 0.93% (micro average: 82.69%)	84.08%	97.56%
K-Nearest Neighbour	72.52% (micro average: 72.52%)	85.84%	79.60%

Pada Tabel 4 tersebut dapat dilihat bahwa algoritma *Decision Tree* merupakan algoritma yang paling baik dibanding dengan algoritma ML lain untuk mengenali kemiskinan multidimensi di Provinsi Nusa Tenggara Timur yang diuji pada penelitian ini. Tingkat akurasi algoritma *Decision Tree* untuk mengklasifikasikan kemiskinan multidimensi lebih tinggi dari banding yang lain, begitu juga nilai *recall* pada algoritma ini menghasilkan nilai tertinggi dibanding yang lain.

Berdasarkan pengujian tersebut maka dilakukan pembentukan model kemiskinan multidimensi untuk menentukan klasifikasi kemiskinan multidimensi di Provinsi Nusa Tenggara Timur menggunakan algoritma *Decision Tree*. Hasil pembentukan model tersebut menunjukkan bahwa dimensi kesehatan pada indikator penolong kelahiran anak memiliki *gain* yang tertinggi. Dengan kata lain, indikator penolong persalinan dalam dimensi kesehatan ini menjadi indikator awal dalam pengklasifikasian multidimensional kemiskinan di Provinsi Nusa Tenggara Timur. Secara lengkap visualisasi model algoritma *Decision Tree* dapat dilihat pada Gambar 1.

Path Decision Tree

```

Penolong Persalinan = Bukan Nakes
| Rumah Lavak = Lavak: Tidak Miskin
| Rumah Lavak = Tidak
| Pendidikan Dasar = Ada yang belum
| | Kepemilikan Aset = Memiliki
| | | Kepemilikan Jamkes = Punya: Tidak Miskin
| | | Kepemilikan Jamkes = Tidak
| | | Kepemilikan Tabungan = Punya: Tidak Miskin
| | | Kepemilikan Tabungan = Tidak
| | | Imunisasi/Vaksi Balita = Belum: Miskin
| | | Imunisasi/Vaksi Balita = Sudah
| | | Melek Huruf = Ada yang belum
| | | | Pekerjaan KRT = Primer: Miskin
| | | | Pekerjaan KRT = Tersier: Tidak Miskin
| | | | Pekerjaan KRT = Tidak Bekerja: Tidak Miskin
| | | | Melek Huruf = Sudah Semua: Tidak Miskin
| | | | | Kepemilikan Aset = Tidak Memiliki
| | | | | Pekerjaan KRT = Primer
| | | | | | Kepemilikan Jamkes = Punya: Tidak Miskin
| | | | | | Kepemilikan Jamkes = Tidak: Miskin
| | | | | Pekerjaan KRT = Sekunder: Miskin
| | | | | Pekerjaan KRT = Tersier: Tidak Miskin
| | | | | Pekerjaan KRT = Tidak Bekerja
| | | | | | Imunisasi/Vaksi Balita = Belum: Tidak Miskin
| | | | | | Imunisasi/Vaksi Balita = Sudah: Miskin
| | | | | Pendidikan Dasar = Sudah Semua: Tidak Miskin
| Pendidikan Dasar = Sudah Semua: Tidak Miskin
Penolong Persalinan = Nakes
| Pendidikan Dasar = Ada yang belum
| | Kepemilikan Tabungan = Punya: Tidak Miskin
| | Kepemilikan Tabungan = Tidak
| | | Imunisasi/Vaksi Balita = Belum
| | | | Pekerjaan KRT = Primer
| | | | | Kepemilikan Aset = Memiliki: Tidak Miskin
| | | | | Kepemilikan Aset = Tidak Memiliki
| | | | | Melek Huruf = Ada yang belum
| | | | | | Kepemilikan Jamkes = Punya: Tidak Miskin
| | | | | | Kepemilikan Jamkes = Tidak: Miskin
| | | | | Melek Huruf = Sudah Semua: Miskin
| | | | | Pekerjaan KRT = Tersier: Tidak Miskin
| | | | | Pekerjaan KRT = Tidak Bekerja: Miskin
| | | | | Imunisasi/Vaksi Balita = Sudah: Tidak Miskin
| Pendidikan Dasar = Sudah Semua: Tidak Miskin
  
```

Gambar 1. Visualisasi *Path Description*
Decision Tree

Gambar 1 di atas juga memperlihatkan bahwa terdapat delapan jalur ciri pembentuk klasifikasi kemiskinan multidimensi di Provinsi Nusa Tenggara Timur, diantaranya adalah:

1. Penolong Persalinan -> Rumah Tinggal -> Pendidikan dasar -> Kepemilikan aset -> Jaminan Kesehatan -> Tabungan -> Imunisasi/Vaksin
2. Penolong Persalinan -> Rumah Tinggal -> Pendidikan dasar -> Kepemilikan aset -> Jaminan

- Kesehatan -> Tabungan -> Imunisasi/Vaksin -> Melek huruf
3. Penolong Persalinan -> Rumah Tinggal -> Pendidikan dasar -> Kepemilikan aset -> Pekerjaan KRT -> Jaminan Kesehatan
 4. Penolong Persalinan -> Rumah Tinggal -> Pendidikan dasar -> Kepemilikan aset -> Pekerjaan KRT
 5. Penolong Persalinan -> Rumah Tinggal -> Pendidikan dasar -> Kepemilikan aset -> Pekerjaan KRT -> Imunisasi/Vaksin
 6. Penolong persalinan -> Pendidikan dasar -> Kepemilikan Tabungan -> Imunisasi/Vaksin -> Pekerjaan KRT
 7. Penolong persalinan -> Pendidikan dasar -> Kepemilikan Tabungan -> Imunisasi/Vaksin -> Pekerjaan KRT -> Kepemilikan aset -> Melek huruf -> Jaminan Kesehatan
 8. Penolong persalinan -> Pendidikan dasar -> Kepemilikan Tabungan -> Imunisasi/Vaksin -> Pekerjaan KRT -> Kepemilikan aset -> Melek huruf.

Dari hasil delapan jalur pembentuk ciri klasifikasi kemiskinan multidimensi tersebut dapat dilihat bahwa selain dimensi kesehatan khususnya indikator penolong persalinan pada wanita pernah melahirkan, indikator pendidikan dasar pada dimensi pendidikan menjadi indikator penting yang menentukan kemiskinan multidimensi di Provinsi Nusa Tenggara Timur. Hal ini dilihat dari indikator ini menduduki posisi *node*

kedua atau ketiga setelah indikator penolong persalinan serta indikator ini selalu ada pada kedelapan jalur tersebut.

4. Simpulan dan Saran

Masalah kemiskinan merupakan persoalan klasik yang terjadi di Indonesia (Ferezagia, 2018). Adanya pandemi Covid-19 telah memperburuk kondisi kemiskinan yang terjadi, oleh karena itu perlu penanganan yang komprehensif dari berbagai bidang dengan data yang akurat. Dari hasil kajian ini dapat ditunjukkan bahwa penggunaan *machine learning* dengan algoritma *Decision Tree* terbukti dapat dimanfaatkan untuk mengklasifikasikan kemiskinan multidimensi di Provinsi Nusa Tenggara Timur dengan tingkat akurasi, presisi dan recall yang cukup tinggi.

Model klasifikasi yang dibentuk melalui algoritma *Decision Tree* menunjukkan bahwa kesehatan dan pendidikan memberi pengaruh tertinggi atas terjadinya kemiskinan multidimensi di Provinsi Nusa Tenggara Timur. Karena itu, penelitian ini merekomendasikan agar kedua indikator tersebut perlu mendapatkan prioritas utama dalam upaya pemerintah menanggulangi kemiskinan. Hasil ini sejalan dengan penelitian Rahajeng dkk, (2021) yang menyimpulkan bahwa faktor pendidikan dan kesehatan memiliki hubungan erat dengan kemiskinan. Oleh karena itu pemerintah perlu lebih memfokuskan pada kedua dimensi ini agar penanganan kemiskinan dapat

dilakukan lebih efektif. Beberapa hal yang dapat dilakukan diantaranya memperbanyak tenaga medis khususnya untuk penolong kelahiran dan memperluas jangkuan pelayanannya agar semua masyarakat dapat mengakses layanan tersebut dengan mudah dan murah. Pada dimensi pendidikan, pemerintah perlu mengencangkan kembali program wajib belajar 9 tahun untuk memenuhi kebutuhan pendidikan dasar. Bagi penduduk usia sekolah harus diberikan akses seluas-luasnya terhadap fasilitas sekolah dan kemudahan biaya pendidikan, sedangkan bagi penduduk yang sudah lewat usia pendidikan namun belum menyelesaikan pendidikan dasar harus diberikan fasilitas Kejar Paket A dan Kejar Paket B. Kesadaran akan tingginya pengaruh pendidikan dan kesehatan diharapkan dapat mengatasi permasalahan kemiskinan multidimensi di Provinsi Nusa Tenggara Timur ini.

Penelitian ini masih terbatas baik penggunaan series data maupun indikator yang digunakan. Pengembangan kedepan tentang pengklasifikasian kemiskinan mutidimensi di Provinsi Nusa Tenggara Timur ini dapat dilakukan dengan memperpanjang series data yang digunakan. Prinsip dasar ML adalah semakin menggunakan banyak data *training* akan meningkatkan akurasi dan presisi algoritma tersebut. Selain itu pengembangan penelitian selanjutnya juga dapat dilakukan dengan memperkaya dimensi atau indikator

dalam menggambarkan kemiskinan di Provinsi Nusa Tenggara Timur.

Daftar Pustaka

- Alfista, M. (2019). *Analisis Tingkat Kemiskinan di Indonesia tahun 2012–2017*, Universitas Muhammadiyah Surakarta
- Alkire, S., & Santos, M. E. (2013). *A multidimensional approach: poverty measurement & beyond. Social indicators research*, 112(2), 239-257.
- Ayudhitama, A. P., & Pujiyanto, U. (2020). *Analisa 4 Algoritma Dalam Klasifikasi Liver Menggunakan Rapidminer*. *Jurnal Informatika Polinema*, 6(2), 1-9.
- BPS, (2022). *Berita Resmi Statistik: Profil Kemiskinan di Provinsi Nusa Tenggara Timur September 2021*. Kupang, BPS Provinsi Nusa Tenggara Timur.
- Budiantoro, S., Fanggidae, V., Saputra, W., Maftuchan, A., & Artha, D. R. P. (2013). *Multidimensional Poverty Index (MPI): Konsep dan Pengukurannya di Indonesia*.
- Djunawan, A. (2019). *Benarkah Subsidi Jaminan Kesehatan Meningkatkan Pemanfaatan Pelayanan Kesehatan Primer oleh Penduduk Miskin Perkotaan*. *Jurnal Kebijakan Kesehatan Indonesia: JKKI*, 8(1), 18-24.
- Fadlillah, N., Dewi, A.S., & Sukiman. (2016). *Analisis Pengaruh Pendapatan Per Kapita, Tingkat Pengangguran, IPM dan Pertumbuhan Penduduk terhadap Kemiskinan di Jawa Tengah Tahun*

- 2009-2013. *Jurnal Pembangunan Ekonomi Wilayah Eko-Regional*, 11(1)
- Ferezagia, D. V. (2018). *Analisis tingkat kemiskinan di Indonesia*. *Jurnal Sosial Humaniora Terapan*, 1(1)
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.
- Jordan, M. I., & Mitchell, T. M. (2015). *Machine learning: Trends, perspectives, and prospects*. *Science*, 349(6245), 255-260.
- Kaunang, F. J. (2019). *Penerapan Algoritma J48 Decision Tree Untuk Analisis Tingkat Kemiskinan di Indonesia*. *Cogito Smart Journal*, 4(2), 348-357.
- Rahajeng, A. S., Muljaningsih, S., & Asmara, K. (2021). *Pengaruh Pertumbuhan Ekonomi, Investasi, Pendidikan, dan Kesehatan terhadap Kemiskinan di Provinsi Jawa Timur*. *Jurnal Syntax Admiration*, 2(5), 774-784.
- Sani, N. S., Rahman, M. A., Bakar, A. A., Sahran, S., & Sarim, H. M. (2018). *Machine learning approach for bottom 40 percent households (B40) poverty classification*. *Int. J. Adv. Sci. Eng. Inf. Technol*, 8(4-2), 1698.
- Sari, A. C. D. M., & Purwanti, E. Y. (2012). *Pengaruh Kepemilikan Aset, Pendidikan, Pekerjaan dan Jumlah Tanggungan terhadap Kemiskinan Rumah Tangga di Kecamatan Bonang Kabupaten Demak*, Fakultas Ekonomika dan Bisnis.
- Sartika, D., & Sensuse, D. I. (2017). *Perbandingan algoritma klasifikasi Naïve Bayes, Nearest Neighbour, dan Decision Tree pada studi kasus pengambilan keputusan pemilihan pola pakaian*. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 3(2), 151-161.
- Setyadi, S., & Indriyani, L. (2021). *Dampak Pandemi Covid-19 Terhadap Peningkatan Resiko Kemiskinan di Indonesia*. *Pareto: Jurnal Ekonomi Dan Kebijakan Publik*, 4(1), 53-66. doi:10.32663/pareto.v4i1.1891
- Sumargo, B., & Simanjuntak, N. M. M. (2019). *Deprivasi Utama Kemiskinan Multidimensi Antarprovinsi di Indonesia*. *Jurnal Ekonomi dan Pembangunan Indonesia*, 19(2), 160-172.
- Tarigan, H., Sinaga, J. H., & Rachmawati, R. R. (2020). *Dampak pandemi Covid-19 terhadap kemiskinan di Indonesia*. *Pusat Sosial Ekonomi dan Kebijakan Pertanian*, 3, 457-479.
- Watratan, A. F., & Moeis, D. (2020). *Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia*. *Journal of Applied Computer Science and Technology*, 1(1), 7-14.
- Yunita, D. (2017). *Perbandingan Algoritma K-Nearest Neighbor dan Decision Tree untuk Penentuan Risiko Kredit Kepemilikan Mobil*. *Jurnal Informatika Universitas Pamulang*, 2(2), 103-107.