

Analisa Prediksi Kesejahteraan Masyarakat Nelayan Lombok Timur Menggunakan Algoritma Random Forest

Arnita Sandi^{1*}, Kusrini², Kusnawi³

^{1,2,3}Program Studi Magister Teknik Informatika, Universitas Amikom Yogyakarta

*arnita_s@students.amikom.ac.id

Abstrak

Kehidupan perekonomian masyarakat di pesisir pantai khususnya nelayan sangat bergantung dengan sumberdaya alam yang ada disekitar, contohnya adalah sumberdaya laut, dimana masih mendapatkan posisi teratas dalam kelangsungan hidup masyarakat nelayan yang banyak dimanfaatkan dan juga termasuk kedalam sumber daya alam yang dapat diperbaharui. Salah satu contoh yang dijadikan sebagai bahan penelitian ini adalah masyarakat nelayan di Lombok Timur Nusa Tenggara barat. Masyarakat Nelayan dapat diartikan sebagai kumpulan dari masyarakat yang Sebagian besar anggotanya bermatapencaharian sebagai nelayan. Karakteristik hidup masyarakat ini berbeda dengan masyarakat pada umumnya. Faktor alam banyak mempengaruhi kehidupan mereka, dari gaya hidup bahkan sampai dengan tingkat perekonomian dan kesejahteraan mereka berbeda dengan masyarakat lainnya. Tujuan dari penelitian ini adalah untuk memprediksi tingkat kesejahteraan masyarakat nelayan di Lombok Timur Nusa Tenggara Barat dengan menggunakan metode klasifikasi dan algoritma Random Forest. Dataset yang digunakan bersifat private data, data diambil dari aplikasi nelayan. Pengolahan data dilakukan untuk mendapatkan hasil atau performa dari algoritma tersebut sebagai hasil terbaik dalam memprediksi. Dari dataset yang ada kita menggunakan lima variabel pendukung diantaranya, Pendidikan, anggota keluarga, sumur (berkaitan dengan air bersih), pekerjaan dan perumahan. Dengan hasil atau target dari pengolahan data ini adalah tingkat kesejahteraan masyarakat nelayan dengan status sejahtera dan tidak sejahtera. Hasil akhir dari penelitian ini dilihat menggunakan Confusion Matrix, dimana hasil akhirnya adalah nilai akurasi. Random Forest memiliki nilai akurasi tertinggi dengan nilai 93, 37% dan nilai AUC 0,735%.

Kata kunci : Klasifikasi, Pediksi, Random Forest

Abstract

The economic life of the people on the coast, especially fishermen, is very dependent on the natural resources that are around, for example, marine resources, which still get the top position in the survival of fishing communities which are widely used and are also included as renewable natural resources. One example used as material for this research is the fishing community in East Lombok, West Nusa Tenggara. The Fishermen's Community can be interpreted as a group of people whose main livelihood is fishermen. The characteristics of the life of this community are different from society in general. Natural factors influence their lives a lot, from their lifestyle to the level of their economy and welfare, which is different from other communities. The purpose of this study is to predict the level of welfare of fishing communities in East Lombok, West Nusa Tenggara by using the classification method and the Random Forest algorithm. The dataset used is private data, the data is taken from fishing applications. Data processing is done to get the result or performance of the algorithm as the best result in predicting. From the existing dataset we use five supporting variables including, Education, family members, wells (related to clean water), employment and housing. The results or targets of this data processing are the level of welfare of fishing communities with prosperous and non-prosperous statuses. The final results of this study are seen using the Confusion Matrix, where the end result is the accuracy value. Random Forest has the highest accuracy value with a value of 93.37% and an AUC value of 0.735%..

Keywords : Classification, Pediction, Random Forest.

1. Pendahuluan

Kehidupan sosial masyarakat nelayan sangatlah beragam dan kompleks dimana kita dapat melihat berbagai macam bentuk aktivitas kelautan yang terjadi di lingkungan masyarakat. Adanya keterikatan antara kebudayaan kehidupan masyarakat dengan budaya maritim membuat sektor perekonomian masyarakat nelayan atau masyarakat maritim lebih homogen. Ini yang menyebabkan perlu adanya solusi dalam pengelolaan dan pemanfaatan sumberdaya alam yang kontinu agar terwujud kesejahteraan masyarakat secara bersama[1].

Karakteristik kehidupan masyarakat pesisir berbeda dengan masyarakat umumnya, sehingga kehidupan perekonomian mereka berbeda. Masyarakat pesisir pantai biasanya hanya bermatapencaharian sebagai nelayan atau petani tambak, sehingga dapat dikatakan sebagai masyarakat yang jauh dari kata sejahtera. Pemerataan pembangunan demi kesejahteraan masyarakat harus dilakukan secara adil dan merata, karena pemerataan pembangunan di berbagai sektor dapat berhubungan dengan pertumbuhan ekonomi masyarakat[2].

Lombok Timur adalah salah satu kota/kabupaten yang ada di provinsi Nusa Tenggara Barat yang memiliki Panjang pantai 220 km dan luas perairan laut kurang lebih sekitar 1.074 km². Dari keadaan ini dapat disimpulkan potensi pengembangan sumberdaya perikanan dan

kelautan di wilayah Lombok Timur cukup tinggi, meskipun pemanfaatan sumberdaya alam tersebut masih tergolong tradisional[3].

Pemanfaatan yang masih bersifat tradisional membuat tingkat kesejahteraan masyarakat nelayan masih dibawah sejahtera atau dapat dikatakan sebagai masyarakat miskin. Tingkat kesejahteraan masyarakat dapat dilihat melalui beberapa faktor diantaranya adalah Pendidikan, pekerjaan, perumahan, anggota keluarga, dan sumur (air bersih). Air bersih masuk kedalam faktor pendukung kesejahteraan masyarakat, karena kualitas air berkaitan erat dengan kehidupan sehari-hari masyarakat, terutama masyarakat nelayan. Keterkaitan kualitas air bersih ini juga pernah dilakukan penelitian oleh Muhammad Malik Mutoffar dan rekan-rekan yang dilakukan di wilayah DKI Jakarta. Dimana tingkat kebutuhan air bersih meningkat hampir 70% berasal dari air sumur. Penelitian ini menghasilkan nilai presisi sebesar 0,823 dan dapat memprediksi air sebesar 82% dari 83% data yang diklasifikasikan menggunakan algoritma Random Forest. Hasil diperoleh dari jumlah data training sebesar 214 dan 53 data testing. Algoritma ini digunakan sebagai metode klasifikasi dan regresi[4].

Algoritma Radom Forest adalah sebuah metode yang berarti hutan acak yang merupakan kombinasi dari pohon-pohon keputusan, dimana setiap pohon memiliki nilai-nilai vektor yang

secara acak atau random di ambil dengan nilai distribusi yang sama untuk semua pohon, metode ini juga banyak digunakan karena memiliki tingkat error yang rendah dalam memproses pemilihan fitur[5].

Penelitian kali ini menggunakan Random Forest yang bertujuan untuk memprediksi tingkat kesejahteraan masyarakat nelayan di Lombok Timur Nusa Tenggara Barat, dan juga untuk mengetahui performance algoritma random forest dalam memprediksi sebuah dataset. Metode ini digunakan untuk menganalisis dan memprediksi data dengan hasil akhir sebuah keputusan sejahtera dan tidak sejahtera untuk masyarakat nelayan di Lombok Timur

2. Tinjauan Pustaka

2.1. Penelitian Terkait

Adapun penelitian sebelumnya yang menggunakan algoritma Random Forest sebagai metode dalam pengolahan data sebagai pendukung penelitian kali ini, antara lain:

1. Eka Christy dan rekan-rekan yang berjudul "Analisis Klasifikasi Status Bekerja Penduduk Daerah Istimewa Yogyakarta Menggunakan Metode Random Forest". Pada penelitian ini dilakukan pemetaan data dari setiap himpunan variable untuk targetnya. Dari penelitian tersebut dihasilkan nilai akurasi sebesar 80,13% dengan variable yang digunakan adalah penyakit, Pendidikan, jenis kelamin dan sertifikat. Penelitian

ini menggunakan algoritma Random Forest menghasilkan nilai akurasi yang cukup tinggi yang dapat digunakan untuk menganalisis status bekerja penduduk selain itu juga penelitian ini bermanfaat sebagai pengembangan ilmu pengetahuan tentang klasifikasi data menggunakan algoritma Random Forest[6].

2. Marchell Rianto dan Roni Yunis yang berjudul "Analisis Runtun Waktu Untuk Memprediksi Jumlah Mahasiswa Baru dengan Model Random Forest", penelitian ini dilakukan untuk menyaring mahasiswa yang akan masuk sebagai mahasiswa baruseseuai dengan ketentuan atau kriteria yang berlaku di perguruan tinggi tersebut. Tujuan dari penelitian ini adalah untuk memprediksi jumlah mahasiswa baru. Algoritma random forest digunakan karena dianggap memiliki performance yang sangat baik dalam memecahkan masalah. Nilai akurasi yang didapat untuk penelitian ini adalah 99,8% dan nilai MSE serta MAE sebesar 0,02%. Hasil dari algoritma ini sangat tinggi pada penelitian yang dilakukan sehingga dapat memprediksi jumlah mahasiswa di tahun depan dengan hasil penurunan jumlah mahasiswa. Dengan demikian kinerja Random Forest dapat membantu dalam pengambilan keputusan[7].

3. Budi Prasojito dan Emy Haryanti dalam judul "Analisa Prediksi Kelayakan Pemberian Kredit Pinjaman dengan Metode Random Forest". Penelitian ini menggunakan metode Random Forest untuk memprediksi kelayakan dan tidak

layaknya seorang nasabah dalam pemberian kredit pinjaman. Hasil akurasi dari algoritma tersebut adalah sebesar 83% dan termasuk ke dalam very good model. Hasil akurasi didapat dari pembagian dataset menjadi 80% sebagai data latih dan 20% sebagai data uji. Penelitian dilakukan untuk menerapkan metode klasifikasi dengan menggunakan algoritma random forest agar dapat menganalisa hasil terbaik pada setiap kreditur[8].

4. Aditya Yudha Perdana dan kawan-kawan dalam penelitiannya yang berjudul “Prediksi Stunting Pada Balita Dengan Algoritma Random Forest”. Tujuan dari penelitian ini adalah untuk memprediksi keadaan stunting pada balita di Kecamatan Pitu Kabupaten Ngawi serta untuk mengembangkan sebuah sistem perangkat lunak berbasis website dan android yang dapat melakukan pengukuran keadaan stunting balita. Hasil akurasi yang didapat dari penelitian ini adalah sebesar 97,89%[9].

5. Penerapan metode datamining dengan model Random Forest dan C4.5 digunakan pada penelitian ini untuk menghasilkan nilai akurasi yang dapat digunakan untuk memprediksi prestasi akademik mahasiswa. Pada penelitian ini dilakukan pengujian menggunakan cross validation dan confusion matrix sehingga menghasilkan nilai akurasi, precision dan recall. Hasil akurasi untuk penelitian ini didapat 92,4% nilai tertinggi dari random forest dibandingkan

dengan decision tree C4.5. Penggalan data dilakukan pada penelitian untuk mendapatkan informasi yang baru dan bermanfaat untuk kepentingan perguruan tinggi tersebut dalam memprediksi prestasi akademik mahasiswa, sehingga dapat menyelesaikan masalah-masalah yang timbul dan berdampak pada hasil belajar mahasiswa[10]

2.2. Landasan Teori

1. Random Forest

Random Forest adalah sebuah metode yang memiliki konsep pembuatan sejumlah besar pohon keputusan dimana semua pohon berkolerasi dan bertindak sebagai model ansambel. Prediksi kelas dan keputusan diletakkan di setiap pohon keputusan dengan dasar hasil maksimum[11]. Adapun beberapa tahapan dalam penggunaan metode random Forest yaitu

- Bootstrap, tahap ini adalah tahap penarikan acak dengan sebuah pemulihan yang berukuran n dari data training
- Random sub-setting, tahap ini adalah tahap penyusunan tree berdasarkan data yang ada dengan dilakukan proses pemisahan terbaik secara acak $m < d$ dengan peubah yang jelas.
- Mengulangi langkah a dan b sebanyak k kali sampai menghasilkan k buah tree yang acak

- Tahap yang terakhir adalah melakukan pendugaan gabungan atas dasar k buah tree tersebut[12].

2. Klasifikasi

Klasifikasi adalah sebuah metode dari datamining yang biasa digunakan untuk suatu bidang pembelajaran dimana proses Analisa data dilakukan sehingga menciptakan model-model untuk menggambarkan kelas-kelas dari data yang sedang dianalisa[13].

3. K-Fold Cross Validation

K-Fold Cross Validation digunakan untuk melakukan pelatihan dan pengujian pada dataset, dimana pembagian dilakukan secara efektif dengan ukuran K sama untuk mendapatkan nilai akurasi. Pada prosesnya nilai k ditentukan sebagai sepuluh partisi dan diulangi sepuluh kali hingga kekakuratan dan kinerja model dapat ditemukan dengan nilai rata-rata sepuluh kali lipat[14].

4. Confusion Matrix

Confusion Matrix adalah sebuah metode yang digunakan untuk melihat hasil kinerja dari metode klasifikasi, dan biasanya menggunakan 4 istilah dari hasil proses klasifikasi yaitu True Positif (TP), True Negatif (TN), False Positif (FP) dan false Negatif (FN)[15].

5. AUC

Area Under The Curve (AUC) adalah daerah yang terdapat di bawah Receiver Operating Characteristic (ROC). Dimana ROC adalah

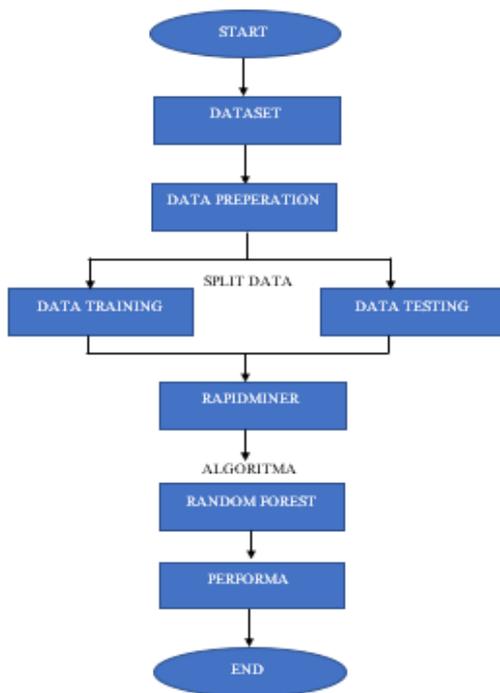
sebuah kurva hasil dari proses antara sensitivitas dan spesifikasi dari berbagai titik potong. Secara teori nilai AUC adalah antara 0 dan 1. Nilai AUC adalah nilai yang memberikan gambaran pengukuran terhadap kesesuaian yang dihasilkan dari algoritma yang digunakan, semakin besar nilai area under curve maka nilai variable yang diteliti untuk memprediksi semakin baik[16].

6. RapidMiner

Rapid miner pertama-tama bernama YALE (Yet Another Learning Environment) dan berubah menjadi Rapidminer pada tahun 2007. Rapidminer adalah sebuah software yang berbasis open-source yang didalamnya terdapat aplikasi pengolahan data dan sebagai mesin data mining yang digunakan untuk pemodelan data, metode visualisasi, transformasi data serta loading data[17]

2.3. Tahapan Penelitian

Adapun alur proses pada penelitian ini dapat terlihat pada gambar 1 dibawah ini :



Gambar 1. Alur proses penelitian

Alur proses penelitian diatas terdiri dari beberapa tahapan yaitu:

- Dimulai dengan dataset, dataset di peroleh dari aplikasi nelayan dalam bentuk .xls. dan .csv
- Tahap selanjutnya adalah data preparation yaitu merubah dataset yang masih mentah menjadi sebuah dataset yang siap untuk dianalisis
- Selanjutnya adalah pembagian data atau split data, data dibagi menjadi data testing dan data training
- Kemudian data yang sudah siap dapat diproses menggunakan software Rapid Miner dan Dengan algoritma random forest, disini

dilakukan proses pemilihan atribut dan target untuk output yang diinginkan

- Hasil dari pemodelan berupa nilai akurasi, precision, recall dan AUC
- Proses selesai sampai mendapatkan hasil atau nilai akurasi tertinggi

3. Metode Penelitian

3.1. Teknik Pengumpulan Data

Pada penelitian data diperoleh dari aplikasi nelayan dan termasuk data private, sedangkan teknik pengumpulan data pada aplikasi nelayan menggunakan teknik observasi, wawancara dan penyebaran angket kepada penduduk.

- Teknik Observasi

Teknik observasi dilakukan dengan melakukan pengamatan langsung ke lokasi penelitian yaitu di kecamatan Labuan Haji, Kecamatan Sakra Timur dan Kecamatan Keruak yang berada di Kabupaten Lombok Timur Nusa Tenggara Barat.

- Teknik Wawancara

Teknik wawancara digunakan untuk mendapatkan data langsung dari penduduk Nelayan di lokasi penelitian

- Angket

Teknik angket ini dilakukan untuk mempermudah saat melakukan wawancara dan juga sebagai alat bukti untuk penginputan data.

Metode penelitian harus dijelaskan secara jelas. Pada bagian ini paling tidak menjelaskan desain penelitian yang digunakan, alasannya

menggunakan desain itu, prosedur penelitian yang dilakukan, populasi dan sampel penelitian atau partisipan penelitian, instrumen penelitian, teknik pengumpulan data, teknik analisis data, dan lokasi penelitian.

3.2. Teknik Analisis Data

Teknik analisis data untuk penelitian kali ini adalah kuantitatif, metode ini digunakan karena sesuai dengan penelitian yang sedang dilakukan. Data yang digunakan adalah data private yang belum pernah digunakan pada penelitian-penelitian sebelumnya. Pada penelitian ini instrument penelitian untuk pengumpulan data adalah observasi langsung ke lokasi, wawancara dan angket.

3.3. Teknik Pengolahan Data

Dalam pengolahan data perangkat lunak yang digunakan adalah Rapidminer dengan metode Klasifikasi dan menggunakan algoritma Random Forest.

3.4. Model Yang diusulkan

Penelitian ini mengusulkan menggunakan metode klasifikasi dengan algoritma Random Forest. Pengujian dilakukan dengan k-fold cross validation dan untuk evaluasi menggunakan confusion matrix dengan hasil akhir nilai akurasi dan nilai AUC serta dalam bentuk grafik.

4. Hasil dan Pembahasan

Proses pengolahan data pada penelitian ini menggunakan software Rapidminer dengan

metode klasifikasi dan menggunakan algoritma Random Forest untuk mengklasifikasikan masyarakat nelayan di Lombok Timur NTB. Jumlah data masyarakat nelayan yang terkumpul adalah 1855 data dengan atribut yang digunakan sebanyak 8 (delapan) atribut, antara lain: Nomor Anggota, Nama, Dusun, Pendidikan, Anggota Keluarga, Sumur, Pekerjaan dan Rumah. Dari delapan atribut tersebut dapat digunakan sebagai komponen penentu keputusan akhir dari keadaan masyarakat nelayan

Tabel 1 Atribut yang digunakan

No	Atribut
1	No Anggota
2	Nama
3	Dusun
4	Pendidikan
5	Anggota Keluarga
6	Sumur
7	Pekerjaan
8	Perumahan

Berikut adalah hasil dari pengolahan data menggunakan rapidminer dengan metode klasifikasi dan algoritma random forest. Pengujian dilakukan dengan K-Fold 10 cross validation.

Tabel 2. Hasil pengolahan data menggunakan Algoritma Random Forest

Cross Validat ion	Hasil Pengolaha Algoritma Random Forest			
	Accurasi %	Precision %	Recall %	AUC %
k-2	93,37	100	93,37	0,597
k-3	93,37	100	93,37	0,701

k-4	93,37	100	93,37	0,694
k-5	93,37	100	93,37	0,735
k-6	93,32	100	93,32	0,663
k-7	93,37	100	93,37	0,694
k-8	93,37	100	93,37	0,667
k-9	93,37	100	93,37	0,723
k-10	93,37	100	93,37	0,726

Dari hasil pengujian menggunakan k-fold cross validation hasil akurasi tertinggi berada di k-5 yaitu 93,37% untuk nilai accurasi dan 0,735 untuk nilai AUC. Dari tabel ini dapat dibuktikan bahwa algoritma random forest sangat baik untuk mengkalsifikasikan data masyarakat nelayan, sehingga nilai yang diperoleh dapat digunakan untuk memprediksi tingkat kesejahteraan masyarakat nelayan di Lombok Timur NTB.

The screenshot shows a 'Multiclass Classification Performance' tool. It displays an accuracy of 93.37% and an AUC of 0.735. Below these metrics is a confusion matrix table:

	true Tidak Sejahtera	true Sejahtera	class precision
pred. Tidak Sejahtera	1732	123	93.37%
pred. Sejahtera	0	0	0.00%
class recall	100.00%	0.00%	

Gambar 2 Hasil Akurasi pada k-Fold 5

Dari tabel pada gambar 2, maka kita dapat melihat True tidak sejahtera sebesar 1732, hasil tersebut adalah hasil klasifikasi yang berarti true positif dimana bahwa faktanya masyarakat nelayan itu tidak sejahtera dan hasil prediksi juga benar. Dan untuk nilai true sejahtera sebesar 123 dan hasil prediksinya adalah salah, sedangkan false negatif sebesar 0 yang artinya antara nilai factual dan hasil klasifikasi adalah salah, bahwa masyarakat nelayan itu tidak sejahtera dan true negatifnya

juga sebesar 0 yang artinya bahwa sebenarnya masyarakat nelayan itu tidak sejahtera tetapi hasil prediksinya salah atau tidak sejahtera. Adapun bentuk tabel dari confusion matrix.

Tabel 3 Tabel Confusion matrix

	True Positif	True Negatif
False positif	1732	123
false negatif	0	0

Adapun perhitungan dari tabel confusion matrix:

- Kita dapat menghitung tingkat error atau kesalahan dengan symbol E dan persamaan berikut

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

$$E = \frac{0 + 123}{0 + 123 + 1732 + 0} \times 100\%$$

$$E = \frac{123}{1855} \times 100\%$$

$$E = 6,63 \%$$

- Setelah itu tingkat akurasi

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

$$= \frac{1732 + 0}{0 + 123 + 1732 + 0} \times 100\%$$

$$= \frac{1732}{1855} \times 100\%$$

$$= 93,37 \%$$

- Dan selanjutnya adalah nilai precision

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

$$= \frac{1732}{0 + 1732} \times 100\%$$

$$= 1 \times 100\% = 100\%$$

- Yang terakhir adalah Recall

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

$$= \frac{1732}{123 + 1732} \times 100\%$$

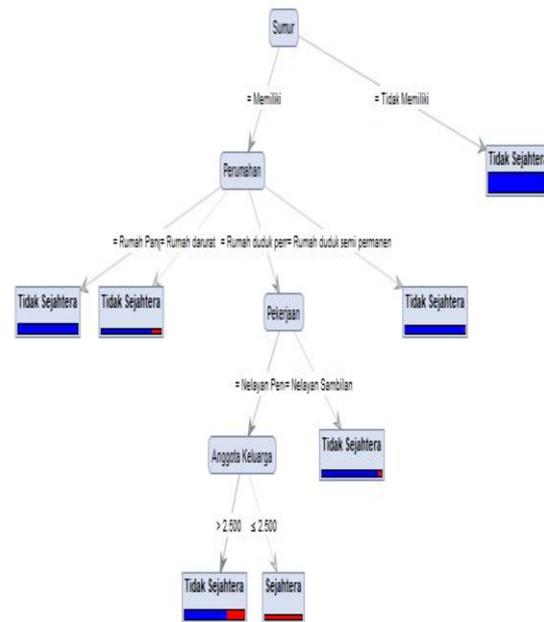
$$= \frac{1732}{1855} \times 100\% = 93,37\%$$

Dari perhitungan tabel confusion matrix tersebut maka dapat kita lihat hasil untuk nilai accurasi sebesar 93,37%, berarti algoritma tersebut telah berhasil mengklasifikasikan tingkat kesejahteraan masyarakat nelayan di Lombok Timur, untuk selanjtnya nilai Precision sebesar 100% ini berarti sistem telah berhasil memberikan informasi yang tepat, sedangkan yang terakhir adalah nilai dari recall sebesar 93,37 yang artinya sistem telah berhasil memberikan informasi yang benar.



Gambar 3 Grafik ROC dan nilai AUC

Grafik ROC menjelaskan bahwa nilai untuk AUC adalah sebesar 0,735 %.



Gambar 4 Random Forest Model

Dari gambar 4 dapat kita lihat masyarakat nelayan tersebut sejahtera atau tidak sejahtera. Masyarakat dapat dikatakan sejahtera jika semua syarat terpenuhi, yaitu: masyarakat memiliki sumur dengan rumah duduk permanen, pekerjaan masyarakat adalah nelayan penuh, berpendidikan, dan anggota keluarga kurang dari atau samadengan dua banyaknya. Jika semua tidak terpenuhi maka masyarakat tersebut dikatakan tidak sejahtera. Dari model random forest tersebut kita dapat melihat bahwa kondisi masyarakat nelayan di Lombok Timur dapat dikatakan tidak sejahtera. Model tersebut adalah hasil dari k-5 validation

5. Kesimpulan

Dari hasil penelitian yang telah dilakukan, metode klasifikasi dengan algoritma random forest dalam memprediksi tingkat kesejahteraan masyarakat nelayan di Lombok Timur NTB, menggunakan 10 kali pengujian dengan k-fold cross validation mendapatkan nilai akurasi sebesar 93,37% dari k-5. Sedangkan untuk nilai AUC diperoleh 0,735%. Dari nilai akurasi yang diperoleh dapat disimpulkan untuk algoritma random forest sangat akurat digunakan untuk menganalisis serta memprediksi tingkat kesejahteraan masyarakat nelayan di Lombok Timur Nusa Tenggara Barat .

6. Daftar Pustaka

- [1] S. Wiranto, "Membangun Kembali Budaya Maritim Indonesia Melalui Kebijakan Kelautan Indonesia dengan Strategi Pertahanan Maritim Indonesia : Perspektif Pertahanan Maritim," pp. 110–126, 2020.
- [2] N. A. Nugroho, "Perbandingan Algoritma Machine Learning dalam Pengklasifikasian Tingkat Kemiskinan di Indonesia Tahun 2021," vol. 6274, pp. 1–12, 2022.
- [3] N. S. Rahmi, "Hubungan Patron-Client Dan Ritual Petik Laut - Studi Kasus Masyarakat Desa Tanjung Luar , Kabupaten Lombok Timur , Nusa Tenggara Barat Oleh Nurbayu Sutiya Rahmi Program Magister Sumber Daya Pantai Universitas Diponegoro."
- [4] M. M. Mutoffar and A. Fadillah, "Klasifikasi Kualitas Air Sumur Menggunakan," vol. 04, no. 02, pp. 138–146, 2022.
- [5] F. F. Veronica Retno Sari, "Perbandingan Prediksi Kualitas Kopi Arabika Dengan Menggunakan Algoritma Sgd, Random Forest Dan Naive Bayes," *EDUMATIC*, vol. 4, no. 2, pp. 1–9, 2020, doi: 10.29408/edumatic.v4i2.2202.
- [6] E. Christy and K. Suryowati, "Analisis Klasifikasi Status Bekerja Penduduk Daerah," vol. 6, no. 1, pp. 69–76, 2021.
- [7] M. Rianto and R. Yunis, "Analisis Runtun Waktu Untuk Memprediksi Jumlah Mahasiswa Baru Dengan Model Random Forest," vol. 23, no. 1, 2021.
- [8] B. Prasajo and E. Haryatmi, "Analisa Prediksi Kelayakan Pemberian Kredit Pinjaman dengan Metode Random," vol. 02, pp. 79–89, 2021.
- [9] A. Y. Perdana, R. Latuconsina, A. Dinimaharawati, and U. Telkom, "Prediksi Stunting Pada Balita Dengan Algoritma Random," vol. 8, no. 5, pp. 6650–6656, 2021.
- [10] S. Linawati and S. Nurdiani, "Prediksi Prestasi Akademik Mahasiswa Menggunakan," vol. VIII, no. 1, pp. 47–52, 2020.
- [11] C. I. Agustyaningrum, W. Gata, R. Nurfalah, and U. Radiah, "Komparasi Algoritma Naive Bayes , Random Forest Dan Svm Untuk Memprediksi Niat," vol. 20, no. 2, 2020.
- [12] M. Azhari, "Perbandingan Akurasi , Recall , dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM, dan Naive Bayes," vol. 5, no. April, pp. 640–651, 2021, doi: 10.30865/mib.v5i2.2937.
- [13] M. Julkarnain, K. R. Ananda, and P. D. Ternak, "Sistem Informasi Pengolahan Data Ternak Unit Pelaksana," vol. 2,

- no. 1, pp. 32–39, 2020.
- [14] M. Asrol, P. Papilo, and F. E. Gunawan, "ScienceDirect Procedia ScienceDirect ScienceDirect Support Vector Machine with K-fold Validation to Improve the Support Vector Machine with K-fold Validation to Improve the Industry ' s Sustainability Performance Classification Industry ' s Sustainability," *Procedia Comput. Sci.*, vol. 179, no. 2020, pp. 854–862, 2021, doi: 10.1016/j.procs.2021.01.074.
- [15] A. Residencia, "Klasifikasi Kelayakan Peserta Pengajuan Kredit Rumah Dengan Algoritma Naïve Bayes Di Perumahan Azzura Residencia," vol. 9, pp. 43–48, 2019.
- [16] L. M. Samsu and W. Amnia, "Implementasi Algoritma Naive Bayes Untuk Klasifikasi Penerima Beasiswa (Studi Kasus Universitas Hamzanwadi)," vol. 6, no. 1, 2023.
- [17] R. Novitrii and N. Irawati; "Integrasi Metode Neive Bayes Dan Software Rapidminer Dalam Analisis Hasil Usaha Perusahaan Dagang Rika Nofitri , Novica Irawati Sekolah Tinggi Manajemen Infomatika dan Komputer Royal," vol. VI, no. 1, pp. 35–42, 2019.