

KLASIFIKASI KUALITAS AIR SUMUR MENGGUNAKAN ALGORITMA *RANDOM FOREST*

Muhamad Malik Mutoffar¹, Muchammad Naseer², Ariansyah Fadillah³

Program Studi Teknik Informatika^{1,2,3}

Sekolah Tinggi Teknologi Bandung^{1,2,3}

malik@sttbandung.ac.id¹, riansyahfadillah@gmail.com²

Abstrak

Jakarta merupakan ibukota negara Indonesia, yang dimana laju perkembangannya sangat pesat di setiap sektor kehidupan. Oleh karena itu kebutuhan air bersih meningkat hampir di perkiraan 70% berasal dari air tanah (air sumur). Masalah banjir yang merupakan persoalan utama yang sering terjadi, yang mengakibatkan kualitas air menurun, khususnya air sumur dan atas dasar kondisi tersebut air di daerah tertentu menjadi tidak layak untuk dikonsumsi. Kehadiran teknologi dapat menjadikan sebuah solusi dalam menangani pencemaran air dalam menentukan kondisi kualitas air tanah, sehingga masyarakat dapat terbantu untuk menentukan kualitas air yang layak. Klasifikasi dalam menentukan kualitas air tanah menggunakan pembelajaran mesin dengan menghasilkan banyak pohon keputusan. *Random Forest* digunakan sebagai klasifikasi dan regresi. Data yang digunakan sebanyak 267 data yang merupakan hasil Analisa laboratorium air sumur di Provinsi DKI Jakarta, dengan data *training* sebanyak 214 data dan data *testing* sebanyak 53 data. Hasil yang didapatkan dari algoritma tersebut adalah presisi dengan nilai 0.823 dan sensitivitas dengan nilai 0.83. Dalam penelitian ini menghasilkan presisi dan sensitivitas yang cukup baik, model algoritma *Random Forest* dapat memprediksi air sebesar 82% dari 83% data yang dapat diklasifikasikan ke dalam air yang dapat dikonsumsi maupun tidak dapat dikonsumsi. Hasil prediksi terdapat 40 data yang berhasil memprediksi nilai 0 sesuai dengan targetnya dan 4 data yang berhasil memprediksi nilai 1 dengan sesuai targetnya terhadap data *testing*.

Kata kunci : klasifikasi, air sumur, *random forest*

Abstract

Jakarta is the capital city of Indonesia, where the rate of development is very fast in every sector of life. Therefore, the need for clean water increases, almost 70% comes from groundwater (well water). The problem of flooding which is the main problem that often occurs, which results in decreased water quality, especially well water and based on these conditions the water in certain areas becomes unfit for consumption. The presence of technology can make a solution in dealing with water pollution in determining groundwater quality conditions, so that people can be assisted in determining proper water quality. Classification in determining groundwater quality uses machine learning by generating many decision trees. *Random Forest* is used as a classification and regression. The data used is 267 data which is the result of laboratory analysis of well water in DKI Jakarta Province, with 214 data training data and 53 data testing data. The results obtained from the algorithm are precision with a value of 0.823 and sensitivity with a value of 0.83. In this study, it produced quite good precision and sensitivity, the *Random Forest* algorithm model could predict water by 82% of the 83% of the data which could be classified into consumable and non-consumable water. The results of the prediction are 40 data that successfully predict the value 0 according to the target and 4 data that successfully predict the value 1 according to the target for data testing.

Keywords : classification, well water, *random forest*

I. PENDAHULUAN

Jakarta merupakan ibu kota negara Indonesia, yang dimana laju perkembangannya yang cepat diberbagai sektor. Oleh karena itu penggunaan air bersih terus meningkat hampir diperkirakan 70% air berasal dari air tanah (sumur) yang sering digunakan dalam aktivitas sehari-hari [1]. Diantaranya banyak perumahan, apartemen, dan gedung-gedung pencakar langit melakukan kegiatan sehari-harinya menggunakan air dalam jumlah yang besar. Masalah yang sering didapatkan terutama pada wilayah pesisir di Indonesia yaitu, rawan terhadap banjir, pengikisan tanah, dan penyusutan air laut [2].

Banjir merupakan masalah yang menjadi persoalan utama di Jakarta, hampir setiap tahun kita sering dengar dalam berita yaitu banjir yang terjadi di Jakarta dan terbukti pada bulan Januari sampai Februari tahun 2020 telah terjadi enam kali banjir [3]. Kebutuhan air bersih sangat diperlukan bagi masyarakat terutama pemakaian air tanah dalam aktivitas sehari-hari, karena air tanah rentan akan penurunan kualitasnya karena adanya pengaruh air sungai sekitarnya yang telah tercemar oleh limbah rumah tangga maupun limbah industri [4]. Dalam pernyataan tersebut, kondisi air tanah di Jakarta sangat memprihatinkan karena sering terjadinya banjir dan air sungai yang tercemar dikhawatirkan kualitas air tanah tidak layak konsumsi. Sehingga kualitas air tanah perlu diperiksa atau diuji berdasarkan parameter fisika-kimia yang berada di dalam kandungan air tanah antara lain zat besi, zat mangan, pH, dan sebagainya [5].

Hadirnya teknologi sekarang dapat membantu mempercepat hasil proses dari kualitas air yang dapat dikonsumsi maupun tidak. Kualitas air dapat digambarkan dalam bentuk parameter yang terdiri dari isi kandungan air yang ada di dalamnya, dari parameter atau variabel tersebut dapat digunakan dalam memprediksi kondisi air dengan metode klasifikasi data dengan memanfaatkan *machine learning* [6]. Klasifikasi menjadi metode yang cocok dalam kasus ini untuk membedakan kualitas air yang dapat dikonsumsi ataupun tidak karena klasifikasi terbentuk dari sekumpulan data latih yang sebelumnya telah ditentukan kelasnya [7]. Klasifikasi menggunakan metode *Random Forest* yang akan dipakai dalam penelitian.

Pengklasifikasian ini dipakai karena hasil akurasi yang diberikan serta metode ini sering digunakan sehingga menjadi populer [8]. Adapun tujuan yang didapatkan dari penelitian ini yaitu, mengetahui cara implementasi metode *Random Forset* pada klasifikasi kualitas air tanah dan kondisi kualitas air tanah yang berada di Jakarta. Metode ini digunakan untuk melakukan analisis terhadap kondisi kualitas air dan mengetahui klasifikasi data air tanah yang dapat dikonsumsi maupun tidak.

II. TINJAUAN PUSTAKA

Klasifikasi kualitas air yang akan dikembangkan mengacu pada penelitian yang dilakukan oleh Prismahardi, penelitian tersebut membahas bagaimana menganalisis *dataset water potability* dengan menggunakan Teknik *Lucifer Machine Learning* dengan menggunakan 8 model algoritma klasifikasi. Hasil tingkat akurasi tertinggi didapatkan oleh *Random Forest Classifier* dengan 72,81% dan hasil tingkat akurasi terendah didapatkan oleh *Logistic Regression* dengan 51,86% [6].

1. Air Tanah (Air Sumur)

Air adalah sebuah komponen penting dalam kehidupan bagi semua makhluk hidup, hampir 71% permukaan bumi yaitu perairan [9]. Air menjadi peran penting dalam kegiatan sehari-hari untuk memenuhi kebutuhan bagi manusia, hewan maupun tumbuhan. Selain itu, air dalam tubuh manusia juga dapat membantu proses melarut zat kimia karena hampir 80% bagian dalam tubuh manusia berbentuk cairan [10].

Air tanah merupakan air yang terbentuk dari proses air hujan secara alami, kemudian terserap ke dalam tanah melewati proses lapisan batuan [11]. Sumber air tanah yang sering digunakan dalam kehidupan masyarakat sehari-hari adalah air sumur [12].

2. Kualitas Air

Kualitas air yaitu kandungan makhluk hidup atau zat tertentu yang berada di dalamnya serta dapat dinyatakan oleh beberapa parameter fisika, kimia, dan biologi [13]. Kualitas air telah menjadi istilah penting dalam menggambarkan kesesuaian atau kecocokan air dalam penggunaannya seperti kualitas air minum, kualitas air perikanan, dan sebagainya.

Air minum merupakan air yang kualitasnya telah di uji dan memenuhi persyaratan kesehatan baik secara mikrobiologi, kima, fisika, dan radioaktif yang sudah ditentukan sehingga dapat dikonsumsi [14]. Peraturan Menteri Kesehatan Republik Indonesia No.492/MENKES/PER/IV/2010 Tentang Persyaratan Kualitas Air Minum. Berikut tabel yang merupakan parameter wajib kualitas air minum [15].

TABEL I.
PARAMETER KUALITAS AIR MINUM

Jenis Parameter	Parameter	Satuan	Kadar Maksimum yang Diperbolehkan
Mikrobiologi	E. Coli	Jumlah per 100 mL sampel	0
	Total Bakteri Koliform	Jumlah per 100 mL sampel	0
Kimia Anorganik	Arsen	mg/L	0,01
	Fluorida	mg/L	1,5
	Total Kromium	mg/L	0,05
	Kadmium	mg/L	0,003
	Nitrit	mg/L	3
	Nitrat	mg/L	50
	Sianida	mg/L	0,07
	Selenium	mg/L	0,01
Fisik	Bau		Tidak berbau
	Warna	TCU	15
	TDS	mg/L	500
	Kekeruhan	NTU	5
	Rasa		Tidak berasa
	Suhu	°C	Suhu udara \pm 3
Kimia	Almunium	mg/L	0,2
	Besi	mg/L	0,3
	Kesadahan	mg/L	500
	Khlorida	mg/L	250
	Mangan	mg/L	0,4
	pH		6,5-8,5
	Seng	mg/L	3
	Sulfat	mg/L	250

	Tembaga	mg/L	2
	Ammonia	mg/L	1,5

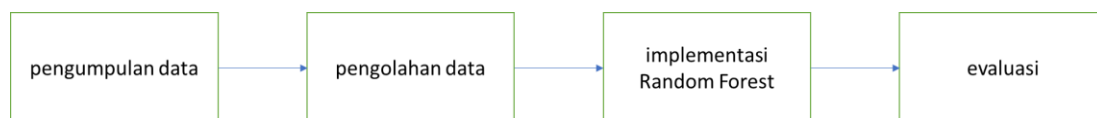
3. Klasifikasi

Kecerdasan buatan menjadi sebuah teknologi yang diterapkan dalam bidang kesehatan. Kemampuan reaktif, dapat mempelajari data historis untuk memberikan prediksi, klasifikasi dalam mengambil keputusan. Sebagian besar aplikasi saat ini mulai dari sistem virtual, chatbots hingga kendaraan self-driving sudah menjadi bagian dari bidang ini, dalam hal ini kecerdasan buatan diterapkan untuk klasifikasi kualitas air sumur [16] [17].

Klasifikasi merupakan suatu pekerjaan untuk mengetahui nilai dari suatu data dan menentukan data tersebut masuk ke dalam kelas tertentu yang telah dibuat, pembangunan modelnya berdasarkan data latih yang ada untuk digunakan mengklasifikasikan pada data yang baru [18] [19]. Cara pengelompokan klasifikasi berdasarkan ciri-ciri yang dimiliki suatu objek dan dapat dilakukan oleh manusia ataupun dengan bantuan teknologi [20]. Salah satunya dengan bantuan teknologi dalam penelitian ini menggunakan *Random Forest*.

III. ANALISIS DAN PERANCANGAN

Penelitian klasifikasi kualitas air dengan menggunakan metode algoritma *Random Forest* yang digunakan terdiri dari 4 tahapan yaitu, pengumpulan data, pengolahan data, implementasi algoritma *Random Forest*, dan evaluasi. Berikut alur metode penelitian dapat terlihat pada Gambar 1.



Gambar 1. Alur Metode Penelitian

1. Pengumpulan Data

Data dalam penelitian ini pengambilan air sumur di Provinsi DKI Jakarta pada tahun 2017 oleh Dinas Lingkungan Hidup Provinsi DKI Jakarta. Dataset yang digunakan adalah kualitas air sumur pada periode ke 2, yang terdiri dari 4272 data dengan 6 atribut, yaitu: (a) periode pengujian, (b) nama lokasi, (c) parameter, (d) satuan, (e) nilai, dan (f) indeks pencemaran.

2. Pengolahan Data

Pada tahap proses pengolahan data, dilakukan menggunakan atribut pada parameter digunakan dan transformasi terhadap data. Atribut yang digunakan yaitu: (a) parameter, (b) nilai, dan (c) indeks pencemaran. Fase pengolahan data selanjutnya adalah transformasi data, dimana data yang berada di dalam parameter di ubah menjadi atribut dan mengganti nilai indeks pencemaran menjadi bernilai 1 jika indeks pencemaran kurang dari sama dengan 1.00 maka kualitas air memenuhi standar dan bernilai 0 jika indeks pencemaran lebih dari 1.00 maka kualitas air tidak memenuhi standar. Jumlah data yang digunakan menjadi 267 data. Contoh transformasi data terdapat pada tabel berikut.

TABEL II.
SEBELUM TRANSFORMASI DATA

parameter	nilai	indeks_pencemaran
Zat Padat Terlarut	254	1.40
Kekeruhan	1.1	1.40
Besi (Fe)	0.13	1.40
Fluorida	0.05	1.40
Total Hardness	93.48	1.40
Chlorida	3.99	1.40
Mangan (Mn)	0.05	1.40
Nitrat	6.25	1.40
Nitrit	0.004	1.40
pH	6.8	1.40
Seng (Zn)	0.08	1.40
Sulfat	40.91	1.40
Senyawa Aktif Biru Metilen	0.03	1.40
Organik (KMnO ₄)	1.64	1.40
Suhu	23.9	1.40
Bakteri Koli	100	1.40

TABEL III.
SETELAH TRANSFORMASI DATA

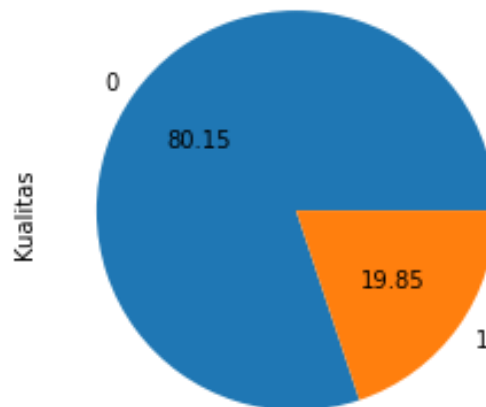
ZatPadatTerlarut	Kekeruhan	Besi	Fluorida	TotalHardness	Chlorida	Mangan	Nitrat	Nitrit	pH	Seng	Sulfat	SenyawaAktifBiruMetilen	Organik	Suhu	BakteriKoli	Kualitas
254	1.1	0.13	0.05	93.48	3.99	0.05	6.25	0.004	6.8	0.08	40.91	0.03	1.64	23.9	100	0

Berikut hasil dari proses pengolahan data yang akan selanjutnya di proses pada tahapan implementasi algoritma *Random Forest*, terdapat beberapa jenis atribut atribut yang akan digunakan terlihat pada tabel 4.

TABEL IV.
JENIS ATRIBUT

No	Nama Atribut
1	Zat Padat Terlarut
2	Kekeruhan
3	Besi
4	Fluorida
5	Total Hardness
6	Chlorida
7	Mangan
8	Nitrat

9	Nitrit
10	pH
11	Seng (Zn)
12	Sulfat
13	Senyawa Aktif Biru Metilen
14	Organik (KMnO ₄)
15	Suhu
16	Bakteri Koli
17	Kualitas

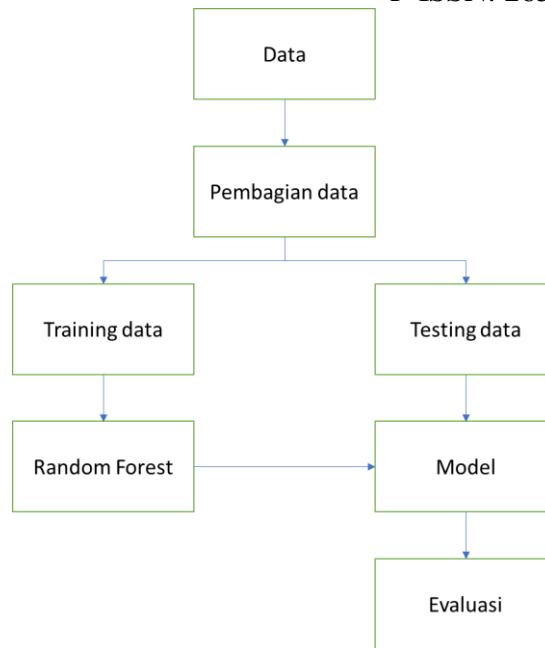


Gambar 3. Jumlah Data Terhadap Kualitas Air

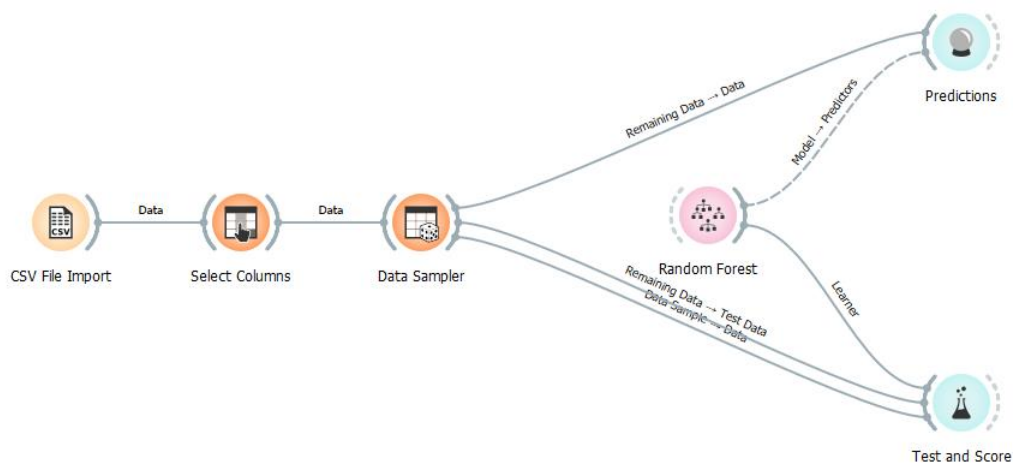
Berdasarkan tampilan dari gambar 3 yang terdiri dari 267 data kualitas air di Jakarta terdapat 80.15% terdiri dari data kualitas bernilai 0 (tidak memenuhi standar) dan 19.85% terdiri dari data kualitas bernilai 1 (memenuhi standar). Adapun perangkat lunak yang digunakan untuk membantu penelitian ini adalah *software Orange 3.32.0*, berikut merupakan pemilihan fitur dan target yang akan digunakan.

3. Implementasi *Random Forest*

Menurut usulan Breiman pendekatan *Random Forest* adalah suatu algoritma pembelajaran mesin dengan menghasilkan banyak pohon keputusan. *Random Forest* merupakan kombinasi dari metode *Bagging* dan *Random Sub Spaces* yang digunakan untuk klasifikasi dan regresi. Ada 3 aspek penting dalam penggunaan metode algoritma ini diantaranya: (1) membangun pohon prediksi dengan melakukan *bootstrap sampling*, (2) memprediksi dengan *predictor* acak pada masing-masing pohon keputusan, dan (3) selanjutnya metode ini dapat melakukan prediksi dengan menggabungkan hasil dari setiap pohon keputusan melalui *majority vote* untuk menentukan klasifikasi atau regresi [21]. Pada Gambar 4 terdapat informasi terkait penerapan implementasi algoritma *Random Forest* untuk klasifikasi kualitas air.



Gambar 4. Tahapan Implementasi *Random Forest*



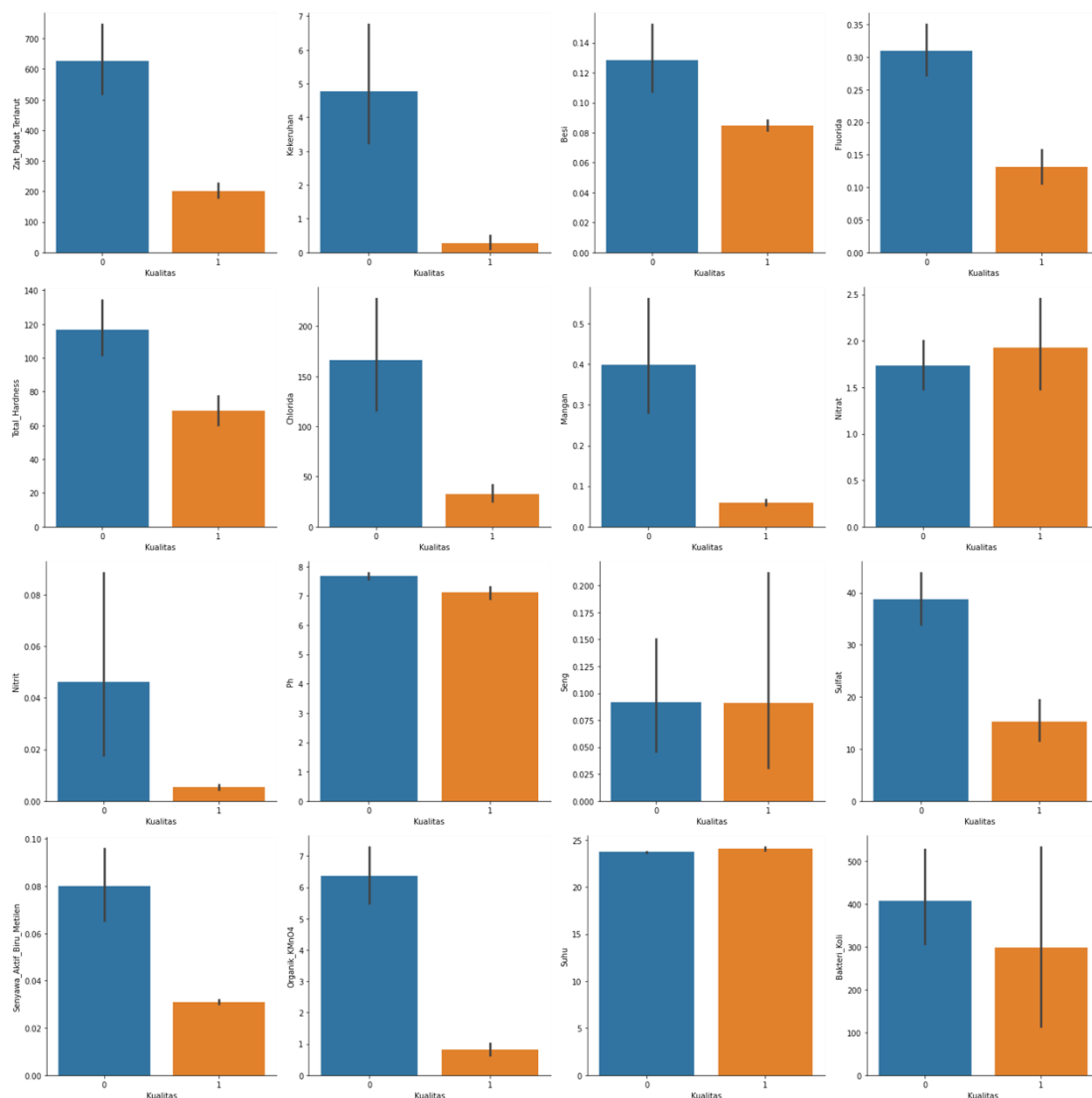
Gambar 5. Tahapan Implementasi *Random Forest Software Orange*

Pada gambar 5 merupakan tahapan analisis data dengan menggunakan algoritma *Random Forest* melalui aplikasi *Orange* yang terdiri *CSV File Import*, *Select Columns*, *Data Sampler*, Algoritma *Random Forest*, *Predictions*, dan *Test and Score*. *CSV File Import* merupakan sebuah alat untuk memasukkan *file* berformat ekstensi *.csv*, berikut tampilan beberapa data dari dataset kualitas-air-sumur- pada gambar 6.

	Zat_Padat_Terlarut	Kekeruhan	Besi	Fluorida	Total_Hardness	Chlorida	Mangan	Nitrat	Nitrit	Ph	Seng	Sulfat	Senyawa_Aktif_Biru_Metilen	Organik_KMnO4	Suhu	Bakteri_Koli	Kualitas
1	254	1.1	0.13	0.05	93.48	30.99	0.05	6.25	0.004	6.8	0.08	40.9	0.03	1.64	23.9	100	0
2	184	0.1	0.08	0.05	76.09	10.33	0.05	0.43	0.003	7.7	0.01	11.13	0.03	2.55	23.9	0	1
3	99	4.3	0.08	0.05	43.48	12.3	0.05	3.61	0.003	6.6	0.02	12.87	0.03	1.99	23.9	0	0
4	127	1.4	0.09	0.05	56.52	0.98	0.09	0.19	0.003	7.5	0.01	8.32	0.03	1.8	23.8	0	1
5	241	0.02	0.08	0.22	97.96	44.01	0.43	2	0.003	6.7	0.01	32.64	0.04	2.43	24.9	100	0
6	100.7	1	0.08	0.05	47.83	2.95	0.05	0.68	0.003	7.5	0.01	8.5	0.03	1.48	23.5	0	1
7	154	0.1	0.08	0.05	47.83	33.95	0.15	2.29	0.003	5.1	0.03	10.85	0.03	2.21	23.6	0	0
8	125	0.1	0.08	0.05	50	14.76	0.05	5.95	0.003	6.2	0.01	11.01	0.03	1.31	23.6	200	0
9	66.7	0.1	0.19	0.05	26.09	15.25	0.77	4.81	0.003	6	0.01	10.96	0.03	1.89	23.4	0	0
10	134	11.2	0.3	0.05	54.35	7.38	0.05	0.85	0.003	6.8	0.01	24.1	0.03	1.61	23.5	200	0
11	196	0.1	0.08	0.05	80.43	26.57	0.05	3.86	0.003	6.7	0.04	21.5	0.03	1.23	23.5	300	0
12	113	0.1	0.08	0.05	41.3	22.14	0.08	2.53	0.004	5.6	0.04	8.33	0.03	1.73	23.8	100	1
13	133	0.1	0.08	0.05	34.78	29.52	0.71	2.58	0.004	4.7	0.07	9.92	0.03	2.11	23.7	0	0
14	118	10.7	0.29	0.05	58.7	3.94	0.05	0.54	0.003	7.1	0.01	9.61	0.03	1.1	23.6	228.8	0
15	285	0.1	0.08	0.05	119.57	41.82	0.05	5.61	0.047	6.8	0.02	33.71	0.03	2.78	23.5	309.92	0
16	106.1	0.1	0.1	0.05	47.83	12.79	0.05	5.4	0.003	6.6	0.02	9.24	0.03	0.91	23.4	618	1
17	72.3	0.1	0.08	0.05	32.61	6.4	0.05	5.19	0.003	6.3	0.01	11.61	0.03	1.7	23.5	200	1
18	155	0.01	0.08	0.37	60.87	21.15	0.05	7.12	0.003	6.4	0.02	9.4	0.03	0.6	23.5	0	0
19	200	0.1	0.08	0.03	76.09	35.91	0.05	1.91	0.003	6.2	0.01	10.6	0.03	2.08	23.5	0	1
20	191	12.1	0.3	0.04	80.43	22.63	0.09	0.69	0.003	6.8	0.01	12.47	0.03	1.45	23.6	100	0

Gambar 6. Data Kualitas Air Sumur Jakarta

Select Columns, yaitu proses pemilihan fitur dan target. Adapun fitur yang digunakan terdiri dari 16 fitur yaitu (1) zat padat terlarut, (2) kekeruhan, (3) besi, (4) fluoride, (5) total hardness, (6) chloride, (7) mangan, (8) nitrat, (9) nitrit, (10) ph, (11) seng, (12) sulfat, (13) senyawa aktif biru metilen, (14) organik KMnO₄, (15) suhu, dan (16) Bakteri Koli. Untuk target dalam penelitian ini menggunakan kualitas, berikut gambar 7 fitur terhadap target yang telah di tentukan.

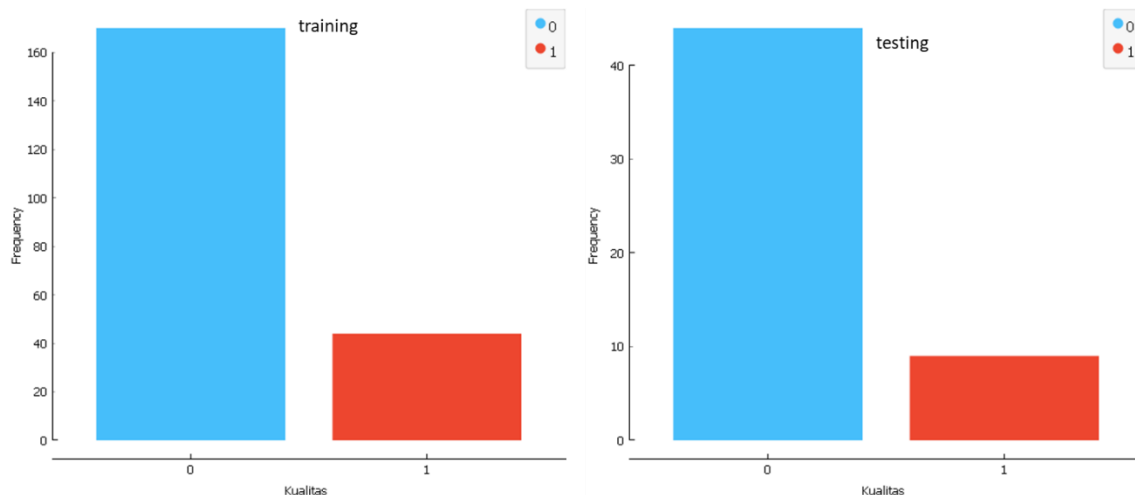


Gambar 7. Tampilan Jumlah Data Fitur Terhadap Target

Data Sampler, yaitu sebuah tahapan dimana pembagian *training data* dan *testing data* terhadap suatu data. Terdapat 267 data yang digunakan dengan pembagian sebesar 80% pada *training data* dan 20% pada *testing data*. *Training data*, yaitu data yang akan dilatih oleh algoritma *Random Forest* sedangkan *testing data*, yaitu data yang digunakan untuk menguji dan mengevaluasi hasil dari model algoritma tersebut.

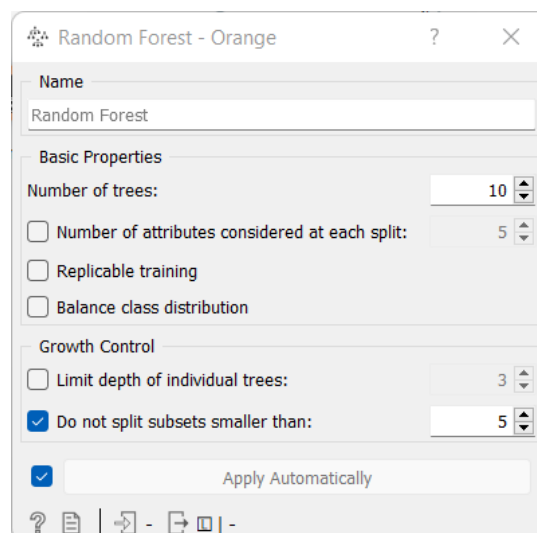
TABEL V.
PEMBAGIAN DATA

Data	Jumlah	Presentase
Training	214	80%
Testing	53	20%
Total	267	100%



Gambar 8. Data Training Dan Testing

Pada Gambar 8 terdapat data *training* dengan jumlah 170 pada data kualitas yang bernilai 0 sedangkan data kualitas bernilai 1 memiliki jumlah 44. Dan data *testing* dengan jumlah 44 pada data kualitas bernilai 0 sedangkan data kualitas bernilai 1 memiliki jumlah 9.

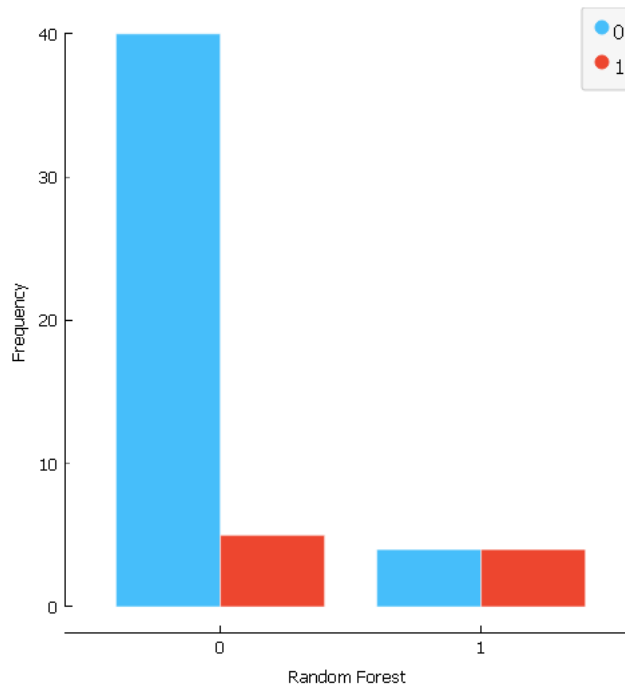


Gambar 9. Pengaturan *Random Forest*

Implementasi algoritma *Random Forest*, yaitu tahapan penerapan data dengan algoritma *Random Forest* dalam penelitian ini menggunakan setelan *default* yang telah disediakan aplikasi *Orange*. Sedangkan tahapan untuk *prediction* dan *test and score* akan dijelaskan dalam tahapan evaluasi.

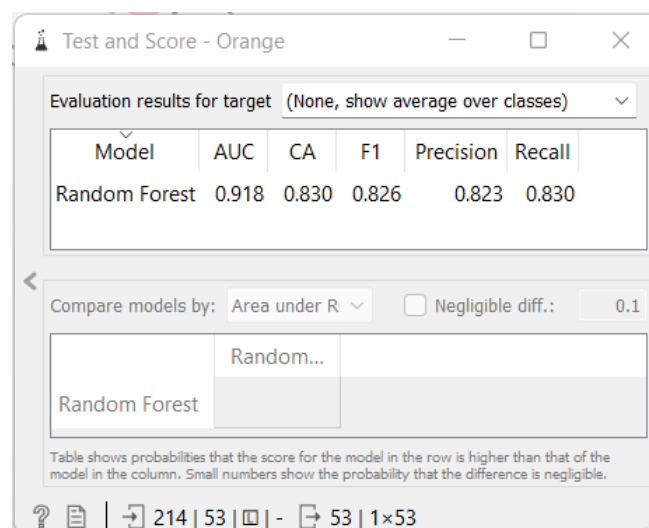
4. Evaluasi

Evaluasi performa *Random Forest* yang akan dilakukan dalam penelitian menggunakan *prediction* dan *test and score* terhadap presisi (*precision*), dan sensitivitas (*recall*) serta menggunakan *prediction* dan *test and score* dari aplikasi Orange untuk mengukur kinerja atau performa model *Random Forest* [25]. *Prediction*, yaitu tahapan untuk melihat hasil prediksi dari data *testing* dengan menggunakan algoritma *Random Forest*. Berikut hasil prediksi terdapat 45 data *testing* dengan target 0 mempunyai 40 data yang berhasil memprediksi target 0 dan 5 data memprediksi target 1, sedangkan 8 data *testing* dengan target 1 mempunyai 4 data yang berhasil memprediksi 1 dan 4 data memprediksi 0.



Gambar 10. Hasil Prediksi *Random Forest* Terhadap Data Testing

Test and Score, yaitu tahapan untuk melihat kinerja atau performa yang dihasilkan algoritma *Random Forest* terhadap data yang telah dianalisis. Berikut hasil yang didapatkan terhadap presisi sebesar 0.823 dan sensitivitas sebesar 0.830.



Gambar 11. Hasil Test dan Score *Random Forest*

IV. KESIMPULAN

Klasifikasi kualitas air sumur di Jakarta dengan menggunakan data sebanyak 267 data, dengan pembagian data terdiri 80% pada *data training* dan 20% pada *data testing* menggunakan algoritma *Random Forest* mendapatkan hasil presisi 0.823 dan sensitivitas sebesar 0.83. Dalam penelitian ini menghasilkan presisi dan sensitivitas yang cukup baik, model algoritma *Random Forest* dapat memprediksi air sebesar 82% dari 83% data yang dapat diklasifikasikan ke dalam air yang dapat dikonsumsi maupun tidak dapat dikonsumsi. Hasil prediksi terdapat 40 data yang berhasil memprediksi nilai 0 sesuai dengan targetnya dan 4 data yang berhasil memprediksi nilai 1 dengan sesuai targetnya terhadap data *testing*.

REFERENSI

- [1] B. C. Matahelumual, "Kajian kondisi air tanah Jakarta tahun 2010," J. Lingkung. dan Bencana Geol., vol. 1, no. 3, pp. 131–149, 2010.
- [2] L. Octonovrilna and I. P. Pudja, "Analisa Perbandingan Anomali Gravitasi Dengan Persebaran Intrusi Air Asin (Studi Kasus Jakarta 2006-2007)," J. Meteorol. dan Geofis., vol. 10, no. 1, 2009.
- [3] K. Puspitasari, "Kapabilitas dan Kepemimpinan Anies Baswedan dalam Penanganan Banjir Jakarta di Detik. com dan Kompas. com," J. Ilmu Komun., vol. 18, no. 2, pp. 221–238, 2020.
- [4] S. D. Nuryana, H. Hidartan, H. F. Yuda, and C. P. Riyandhani, "Penyaringan Unsur-Unsur Logam (Fe, Mn) Air Tanah Dangkal di Kelurahan Jembatan Lima, Tambora, Jakarta Barat," J. Abdi Masy. Indones., vol. 1, no. 3, 2019.
- [5] V. E. Hutagaol and S. Simarmata, "Pengaruh Sumur Resapan Terhadap Kualitas Air Tanah Di Fakultas Teknik Universitas Efarina," J. Tek. Unefa Bunga Rampai Tek. Lingkung. Tek. Inform. dan Tek. Elektro, vol. 3, no. 1, 2017.
- [6] P. A. Riyantoko, T. M. Fahrudin, and K. M. Hindrayani, "Analisis Sederhana Pada Kualitas Air Minum Berdasarkan Akurasi Model Klasifikasi Dengan Menggunakan Lucifer Machine Learning," SENADA, vol. 1, no. 01, pp. 12–18, 2021.
- [7] M. A. Shadiq, "Keoptimalan Naïve Bayes Dalam Klasifikasi," Bandung Ilmu Komput. Univ. Pendidik. Indones., 2009.
- [8] M. Belgiu and L. Drăguț, "Random forest in remote sensing: A review of applications and future directions," ISPRS J. Photogramm. Remote Sens., vol. 114, pp. 24–31, 2016.
- [9] Y. D. Hastuti, E. Nasution, and E. Y. Aritonang, "Perilaku Konsumsi Air Minum Pada Siswa/Siswisma Negeri 3 Medan Tahun 2015," Gizi, Kesehat. Reproduksi dan Epidemiol., vol. 1, no. 3, 2015.
- [10] I. P. T. P. Sari, "Tingkat Pengetahuan Tentang Pentingnyamengkonsumsi Air Mineral Pada Siswa Kelas IV di SD Negeri Keputran A Yogyakarta," J. Pendidik. Jasm. Indones., vol. 10, no. 2, 2014.
- [11] D. Hanifa, I. Sota, and S. S. Siregar, "Penentuan lapisan akuifer air tanah dengan metode geolistrik konfigurasi chlumberger di desa Sungai Jati Kecamatan Mataraman Kabupaten Banjar Kalimantan Selatan," J. Fis. Flux J. Ilm. Fis. FMIPA Univ. Lambung Mangkurat, vol. 13, no. 1, pp. 30–39, 2016.
- [12] A. F. Widiyanto, S. Yuniarno, and K. Kuswanto, "Polusi air tanah akibat limbah industri dan limbah rumah tangga," KEMAS J. Kesehat. Masy., vol. 10, no. 2, pp. 246–254, 2015.
- [13] H. Effendi, "Telaah kualitas air, bagi pengelolaan sumber daya dan lingkungan perairan. Kanisius," Yogyakarta, P, vol. 258, 2003.
- [14] M. N. Mirza, "Hygiene Sanitasi dan jumlah Coliform air minum," KEMAS J. Kesehat. Masy., vol. 9, no. 2, pp. 167–173, 2014.
- [15] Permenkes, "Peraturan Menteri Kesehatan Republik Indonesia No.492/MENKES/PER/IV/2010 Tentang Persyaratan Kualitas Air Minum." 2010.
- [16] M. Malik Mutoffar, P. S. Dey, and D. Das, "Nanointelligence in Functional Food," in Nanotechnology in Functional Foods, New Jersey: Scrivener Publishing LLC, 2022, pp. 329–354. doi: <https://doi.org/10.1002/9781119905059.ch13>.
- [17] I. A. Darmawan, M. F. Randy, I. Yuniarto, M. Malik Mutoffar, and M. T. P. Salis, "Penerapan Data Mining Menggunakan Algoritma Apriori Untuk Menentukan Pola Golongan Penyandang Masalah Kesejahteraan Sosial," Sebatik, vol. 26, no. 1, pp. 223–230, 2022.
- [18] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," J. Media Inform. Budidarma, vol. 4, no. 2, pp. 437–444, 2020.
- [19] I. Yuniarto, A. Kurniawan, and M. Malik Mutoffar, "Comparison of Decision Tree, KNN and Naïve Bayes Methods In Predicting Student Late Graduation In the Informatics Engineering Department, Institute Business XYZ," Adpebi Int. J. Multidiscip. Sci., vol. 1, no. 1, pp. 374–383, 2022.
- [20] A. P. Wibawa, M. G. A. Purnama, M. F. Akbar, and F. A. Dwiyanto, "Metode-metode Klasifikasi," in Prosiding Seminar Ilmu Komputer dan Teknologi Informasi Vol. 2018, vol. 3, no. 1.
- [21] A. Primajaya and B. N. Sari, "Random Forest Algorithm for Prediction of Precipitation," Indones. J. Artif. Intell. Data Min., vol. 1, no. 1, pp. 27–31, 2018.