

Implementasi Metode Decision Tree pada Sistem Prediksi Status Kualitas Produk Minuman A

Abdul Halim Anshor¹⁾, Ahmad Turmudi Zy²⁾

^{1), 2)} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa
Jalan Inspeksi Kalimalang 17530 Kota Bekasi Jawa Barat, Indonesia
Email: abdulhalimanshor@pelitabangsa.ac.id¹⁾, turmudi@pelitabangsa.ac.id²⁾

ABSTRACT

The quality of a beverage product is one of the important items that beverage product entrepreneurs must pay attention to. Good quality beverage products will have an impact on consumers' health. UMKM Buah Sabar is one of the MSMEs located in Bekasi district which produces beverage products A. In the distribution of these beverage products, MSME workers in the delivery section have conditions where the product is out of stock or left over. The reseller must be able to understand whether the status of the remaining product is still of good quality or has been damaged. This is very important to pay attention to because the cooling conditions of each reseller have varying degrees of cold, sometimes also influenced by blackouts and unstable electricity voltage. This condition can cause the quality of product A to decrease. The large number of resellers and products sent will make it difficult for MSME workers to detect the quality of beverage product A. To overcome this problem the researchers found a solution that requires a machine learning method to predict the quality status of product A. In this research, the researchers used the decision tree method to predict the quality status of the product Drink A. The data used are 500 samples of drink product A in the production period from November 2023 to February 2024. The parameters used include temperature, color, taste, aroma, and quality status class of drink product A. The results of this research will show the presentation The accuracy value for the quality of product A is 99.59%, this shows that the decision tree algorithm has very good performance in the process of classifying the quality of beverage product A.

Keywords : MSMEs, Products, Decision Trees, Machine Learning, Resellers.

ABSTRAK

Kualitas suatu produk minuman adalah salah satu item penting yang harus diperhatikan oleh para pengusaha produk minuman. Produk minuman yang berkualitas baik akan berpengaruh bagi kesehatan para konsumen. UMKM Buah Sabar merupakan salah satu UMKM yang berlokasi di kabupaten Bekasi yang memproduksi produk minuman A, Dalam pendistribusian produk minuman tersebut pekerja UMKM bagian pengiriman harus dapat memahami apakah status produk sisa tersebut masih berkualitas baik atau sudah rusak. Hal ini sangat penting diperhatikan karena kondisi pendingin setiap reseller mempunyai tingkat suhu dingin yang bervariasi terkadang juga dipengaruhi faktor mati lampu dan tegangan listrik yang tidak stabil. Kondisi tersebut dapat berpengaruh terhadap kualitas dari produk A menjadi menurun. Banyaknya reseller dan produk yang dikirim akan mempersulit pekerja UMKM dalam mendeteksi kualitas produk minuman A. Untuk mengatasi masalah tersebut peneliti menemukan solusi bahwa diperlukan sebuah metode mesin learning untuk memprediksi status kualitas produk A. Dalam penelitian ini peneliti menggunakan metode decision tree untuk memprediksi status kualitas produk minuman A. Data yang digunakan adalah sampel minuman produk A sebanyak 500 pcs pada periode produksi november 2023 sampai february 2024, adapun parameter yang digunakan antara lain suhu, warna, rasa, aroma dan kelas status kualitas produk minuman A. Hasil penelitian ini akan menunjukkan presentasi nilai akurasi dari kualitas produk A sebesar 99.59%. Ini menunjukkan bahwa algoritma decision tree memiliki kinerja yang sangat baik dalam melakukan proses pengklasifikasian kualitas produk minuman A.

Kata Kunci : UMKM, Produk, Decision Tree, Machine Learning, Reseller.

1. Pendahuluan

Kualitas baik adalah harga mutlak yang harus dimiliki oleh sebuah produk yang akan dipasarkan ke konsumen (*reseller*) karena kualitas produk terutama produk minuman erat kaitannya dengan tingkat kesehatan konsumen yang meminumnya. UMKM Buah sabar adalah salah satu UMKM yang bergerak dibidang produksi produk Minuman kemasan A. Dalam kegiatan sehari-hari produk minuman A. Produk minuman A dibuat dari bahan-bahan alami tanpa bahan pengawet dan dikemas dengan kemasan *cup* (gelas plastik) dan botol plastik. Sehingga dalam penanganan penyimpanan perlu kondisi suhu yang dingin, tidak boleh terpapar matahari dengan intensitas panas yang tinggi dalam waktu yang lama. Produk minuman A yang dihasilkan selanjutnya didistribusikan kepada *reseller* yang terdapat diberbagai wilayah khususnya wilayah Kabupaten Bekasi. Produk minuman A didistribusikan sepekan sekali oleh karyawan (pekerja) UMKM Buah Sabar Ke toko-toko *reseller*. Setiap waktu pendistribusian produk minuman A pekerja bagian pengiriman akan memeriksa atau mengecek status kualitas produk yang dibawa maupun yang masih ada di toko *reseller* apakah kualitas produk minuman A bagus atautakah rusak. Adapun faktor yang dapat mempengaruhi penurunan kualitas produk adalah lemari es atau pendingin *reseller* tidak dingin atau mati (mengalami mati lampu, arus listrik pada lemari pendingin *reseller* tidak stabil, kondisi pemasaran yang lama dengan produk minuman A yang tidak dingin. Rutinitas kegiatan tersebut akan sangat memakan waktu dan menyulitkan bagi pekerja bagian pengiriman dikarenakan jumlah *reseller* dan dan jumlah produk yang sangat banyak, hal ini dapat menimbulkan masalah dimana ada kemungkinan produk minuman A yang rusak bisa terkirim ke toko-toko *reseller*. Untuk mengatasi masalah tersebut peneliti menemukan solusi dengan mengimplementasikan metode *machine learning* yakni *decision tree* untuk memprediksi status kualitas produk minuman A. *Decision tree* merupakan algoritma yang digunakan untuk proses klasifikasi maupun prediksi melalui tahap pembentukan pohon keputusan. Pohon keputusan digunakan untuk prosedur penalaran untuk mendapatkan jawaban dari sebuah masalah (Afdhal, Ariandi, and Rita 2022) (Firdausia Ismi Nurhayati, Nugroho, and Yuniar Purbasari 2021) (Hts 2018). Quinlan (1986) memperkenalkan dalam ID3 induksi decision tree hanya bisa dilakukan pada fitur yang bertipe kategorikal, sedangkan tipe numerik tidak dapat digunakan. Namun quinlan melakukan perbaikan metode ID3 menjadi C4.5 yang dapat menangani fitur tipe numerik, melakukan pemotongan, dan penurunan rule set. Algoritma C4.5 juga dapat menggunakan kriteria gain dalam menentukan fitur untuk pemecahan node dari akar (Lutfi 2018)4). Adapun parameter yang digunakan adalah suhu, warna, rasa, aroma dan kelas status kualitas produk minuman A. Dengan metode ini akan dihasilkan nilai akurasi dari status kualitas produk minuman A yang dapat dijadikan acuan dalam penerapan

standar kualitas produk minuman A di UMKM BUAH SABAR.

Decision Tree

Decision tree (pohon keputusan) adalah salah satu metode *machine learning* yang termasuk dalam kategori supervised learning. Metode ini digunakan untuk membangun model klasifikasi atau regresi. Model yang dihasilkan berbentuk seperti pohon, dimana setiap *node* (simpul) pada pohon tersebut merepresentasikan sebuah keputusan berdasarkan atribut data, dan cabang pada pohon merepresentasikan nilai yang mungkin dimiliki oleh atribut tersebut. Daun (*leaf*) pada pohon *decision tree* merepresentasikan kelas atau nilai output yang diprediksi (Afdhal et al. 2022) (Hts 2018).

Berikut adalah langkah-langkah bagaimana decision tree bekerja dalam proses klasifikasi :

1. Persiapan Data

Kumpulkan data yang akan digunakan untuk membangun model *decision tree*. Data ini harus terdiri dari fitur (atribut) dan label (kelas target).Lakukan preprocessing data, seperti menangani missing *value*, *outliers*, dan normalisasi data jika perlu.

2. Pemilihan Atribut Terbaik

Pada setiap *node* (simpul) di pohon, algoritma *decision tree* perlu memilih atribut terbaik untuk melakukan pembagian data. Pemilihan atribut ini didasarkan pada kriteria tertentu, seperti *Gain Ratio* (C4.5). Mengukur seberapa efektif sebuah atribut dalam memisahkan data ke dalam kelas yang berbeda. Gini Index (CART): Mengukur tingkat ketidakmurnian data pada sebuah *node*. Atribut yang dipilih akan meminimalkan ketidakmurnian pada node anak yang dihasilkan dari pembagian.

3. Pembagian Data

Berdasarkan atribut terbaik yang dipilih, data dibagi menjadi *subset*. Setiap cabang pada pohon *decision tree* mewakili nilai yang mungkin dimiliki oleh atribut tersebut.

4. Pemberhentian (*Stopping Criteria*)

Proses pembagian data akan berhenti ketika salah satu dari kondisi berikut terpenuhi, Semua data pada sebuah node memiliki kelas yang sama (homogen). Tidak ada lagi atribut yang tersedia untuk pembagian data.Kedalaman pohon *decision tree* mencapai batas yang telah ditentukan (biasanya untuk menghindari *overfitting*

5. Penetapan Label (*Leaf Node*)

Node yang tidak dapat dibagi lagi (*leaf node*) diberi label yang sesuai dengan kelas mayoritas data yang ada pada *node* tersebut. Label ini merepresentasikan prediksi model untuk data baru yang masuk ke dalam kategori yang sama.

6. Pemangkasan Pohon (*Optional*)

Setelah pohon decision tree terbentuk, terkadang dilakukan proses pemangkasan (*pruning*) untuk mengurangi kompleksitas model dan mencegah *overfitting*. *Pruning* bisa dilakukan dengan menghilangkan cabang-cabang yang tidak terlalu berkontribusi terhadap akurasi model.

7. Prediksi

Setelah model *decision tree* selesai dibangun, maka dapat digunakan untuk memprediksi kelas dari data baru yang belum pernah dilihat sebelumnya. Data baru akan diturunkan melalui pohon *decision tree* berdasarkan nilai atribut yang dimiliki, hingga mencapai *leaf node*. Label pada *leaf node* tersebut merupakan prediksi kelas untuk data baru tersebut. Berikut dibawah ini adalah persamaan yang berperan dalam pemilihan atribut terbaik untuk pembagian data.

1. Entropy

Konsep entropi digunakan untuk mengukur ketidakmurnian data pada sebuah *node*. Semakin tinggi entropi, maka data tersebut semakin beragam dan sulit diklasifikasikan. Persamaan 1 digunakan untuk menghitung entropi (Negara et al. 2023) (Rewo Ngeo and Latipah 2021) (Wicaksono and Setiadi 2023) (Setiawati, Permana, and Hermawan 2019):

$$Entropy(S) = \sum_{i=1}^m -p(i) \log_2 p(i) \tag{1}$$

Keterangan :

S = Himpunan kasus

m = Banyaknya partisi S

P(i) = Proporsi kelas i (probabilitas yang didapat dari jumlah dibagi total kasus).

2. Information Gain (Gain Informasi)

Information gain mengukur seberapa banyak informasi yang diperoleh dari pembagian data berdasarkan sebuah atribut tertentu. Semakin tinggi information gain, maka atribut tersebut semakin efektif dalam memisahkan data ke dalam kelas yang berbeda. Persamaan 2 digunakan untuk menghitung information gain (Negara et al. 2023) (Rewo Ngeo and Latipah 2021) (Wicaksono and Setiadi 2023) (Setiawati et al. 2019) :

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \tag{2}$$

Keterangan :

S = Himpunan kasus

A = Atribut

n=Jumlah partisi atribut A

|S_i|= Jumlah kasus pada partisi ke i

|S|= Jumlah kasus dalam S

3. Gini Index (Indeks Gini)

Gini index juga merupakan ukuran ketidakmurnian data pada sebuah *node*. Gini index dihitung dengan menjumlahkan probabilitas kuadrat dari data yang salah diklasifikasikan ke dalam setiap kelas. Persamaan 3 digunakan untuk menghitung Gini index adalah (Wicaksono and Setiadi 2023) :

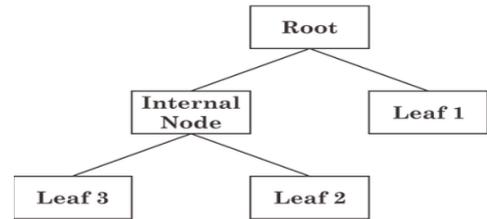
$$Gini(S) = 1 - \sum (p_i)^2 \tag{3}$$

Keterangan:

S: kumpulan data pada sebuah *node*

p_i: proporsi data yang termasuk dalam kelas ke-i

Model pengklasifikasian dengan metode *decision tree* dapat diilustrasikan seperti diperlihatkan pada gambar 1 dibawah ini (Rewo Ngeo and Latipah 2021) :



Gambar 1 Decision Tree

Dalam melakukan penelitian ini peneliti mengacu pada penelitian sebelumnya yang menggunakan metode serupa. Penelitian yang dilakukan oleh Firdausia Ismi Nurhayati dkk, hasil penelitiannya tentang memprediksi status gizi balita Dengan melalui beberapa tahapan, diantaranya dimulai dari pengumpulan data, perhitungan data latih untuk memilih atribut akar atau *parent node*, selanjutnya membuat cabang untuk masing-masing kategori, sehingga menghasilkan suatu pohon keputusan yang sesuai. Pohon keputusan yang dibangun memiliki tingkat akurasi sebesar 0,85, nilai *precision* sebesar 0,2982456 serta nilai *recall* didapatkan sebesar 0.3146. Namun hasil tersebut dapat berubah dalam setiap pemilihan data uji dan data latih. Dalam artian data uji dan data latih yang akan digunakan dalam sistem ini mempengaruhi hasil dari pengukuran kinerja algoritma dan hasil uji coba sistem (Firdausia Ismi Nurhayati et al. 2021). Penelitian kedua dilakukan oleh Moch. Lutfi dengan hasil penelitian dibandingkan dengan tiga kriteria *decision tree* yaitu gain ratio, information gain dan gini index. Hasil akurasi menunjukkan bahwa nilai kriteria tertinggi sebesar 72,76% pada kriteria gini index dan nilai terendah pada kriteria gain ratio sebesar 60,67% (Lutfi 2018). Penelitian ketiga dilakukan oleh Muhammad Afdhal dkk, penelitian menggunakan algoritma *decision tree*, Hasil dari pengembangan sistem adalah prediksi penjualan obat di toko obat.. Hasil dari pengujian *black box* adalah semua sistem yang dikembangkan bekerja dengan baik. Hasil uji validitas dengan membandingkan sistem lama dengan sistem baru dengan rapid miner, dengan menggunakan 30 sampel transaksi akurasi sebesar 89% yang berarti sistem tersebut memiliki kinerja yang baik (Afdhal et al. 2022). Penelitian ke empat dilakukan oleh Oky Irnawati dan Nia Nuraeni dengan hasil penelitian pada dataset *e-commerce gender prediction* menunjukkan adanya peningkatan akurasi *Decision Tree* yang dioptimasi *Particle Swarm Optimization (PSO)* sebesar 0,15%, tingkat akurasi yang sebelumnya sebesar 85.22% dengan angka kurva ROC 0.759 menjadi 85,37% dan kurva ROC 0.764 termasuk dalam fair classification (Irnawati and Nuraeni 2022). Dari penelitian yang dilakukan oleh para peneliti sebelumnya dapat disimpulkan metode *decision tree* dapat melakukan klasifikasi dengan sangat baik (Afdhal et al. 2022) (Firdausia Ismi Nurhayati et al. 2021) (Irnawati and Nuraeni 2022) (Negara et al. 2023) (Rahman, Purnomo, and Hijriana 2022) (Rewo Ngeo and

Latipah 2021) (Sahputra et al. 2020) (Wicaksono and Setiadi 2023) . Dengan demikian peneliti menyimpulkan bahwa algoritma *decision tree* dapat dijadikan solusi dalam pengklasifikasian status kualitas produk minuman A pada UMKM Buah Sabar.

2. Pembahasan

2.1. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data dilakukan dengan cara mengambil data historis penjualan sample produk sebanyak 500 pcs pada periode November 2023 sampai dengan februari 2024. Kemudian peneliti melakukan observasi dan wawancara kepada pihak UMKM berkenaan masalah kualitas produk, dari hasil observasi dan wawancara peneliti menemukan beberapa atribut yang dapat dipakai dalam proses klasifikasi Adapun atribut yang digunakan antara lain, kode produk, nama produk, tanggal produksi, tanggal kadaluarsa, suhu, rasa, warna, dan aroma. Data kualitas produk minuman A yang di produksi pada periode November 2023 sampai februari 2024 diperlihatkan oleh gambar 2

A	B	C	D	E	F	G	H
Kode produk	Nama produk	Tanggal produksi	Tanggal kada luarsa	Suhu (OC)	Rasa	Warna	Aroma
A231101	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231102	Teh Tarik	11/2/2023	11/9/2023	6	Manis, Asam	Coklat pucat	Asam
A231103	Teh Tarik	11/2/2023	11/9/2023	6	Tawar	Coklat krem	Sedikit bau teh
A231104	Teh Tarik	11/2/2023	11/9/2023	10	Pahit	Krem	Susu Basah
A231105	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231106	Teh Tarik	11/2/2023	11/9/2023	3	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231107	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231108	Teh Tarik	11/2/2023	11/9/2023	1	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231109	Teh Tarik	11/2/2023	11/9/2023	3	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231110	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231111	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231112	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231113	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231114	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231115	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231116	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231117	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231118	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231119	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231120	Teh Tarik	11/3/2023	11/10/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar
A231121	Teh Tarik	11/3/2023	11/10/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar

Gambar 2 Data Produk Minuman A Periode November 2023 -2024

2.2. Pembersihan Data

Tahapan selanjutnya adalah pembersihan data, Proses pembersihan data bertujuan untuk meningkatkan kualitas data set kualitas produk minuman A sehingga siap digunakan untuk analisis dan pemodelan. Data yang dibersihkan seperti menghapus data yang *double*, menginput data yang *missing value*, merubah data kategorikal dengan rentang *value* menjadi data numerik dalam kasus ini adalah data suhu.

2.3. Preprocessing Data

Pada tahap *preprocssing* adalah proses pengubahan bentuk data yang belum terstruktur menjadi data yang terstruktur sesuai dengan kebutuhan, untuk diproses lebih lanjut. Data yang telah dibersihkan selanjutnya akan dibeikan label . Proses pelabelan data bertujuan untuk memberikan label pada data kualitas minuman teh tarik agar dapat digunakan untuk membangun model *decision tree*. Model ini nantinya dapat digunakan untuk memprediksi kualitas minuman teh tarik berdasarkan parameter yang ada. Pada penelitian ini peneliti melabelkan data dengan membuatnya menjadi dua kelas yaitu kelas produk berstatus kualitas baik dan produk

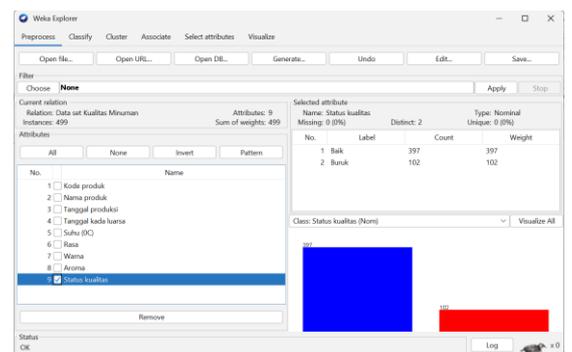
berstatus kualitas buruk. Proses pelabelan data seperti yang diperlihatkan oleh gambar 3 .

A	B	C	D	E	F	G	H	I
Kode produk	Nama produk	Tanggal produksi	Tanggal kada luarsa	Suhu (OC)	Rasa	Warna	Aroma	Status kualitas
A231101	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231102	Teh Tarik	11/2/2023	11/9/2023	6	Manis, Asam	Coklat pucat	Asam	Buruk
A231103	Teh Tarik	11/2/2023	11/9/2023	6	Tawar	Coklat krem	Sedikit bau teh	Buruk
A231104	Teh Tarik	11/2/2023	11/9/2023	10	Pahit	Krem	Susu Basah	Buruk
A231105	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231106	Teh Tarik	11/2/2023	11/9/2023	3	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231107	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231108	Teh Tarik	11/2/2023	11/9/2023	1	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231109	Teh Tarik	11/2/2023	11/9/2023	3	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231110	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231111	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231112	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231113	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231114	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231115	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231116	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231117	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231118	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231119	Teh Tarik	11/2/2023	11/9/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231120	Teh Tarik	11/3/2023	11/10/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik
A231121	Teh Tarik	11/3/2023	11/10/2023	2	Manis, Creamy	Coklat Susu	Teh dan krem segar	Baik

Gambar 3 Proses pelabelan Data

2.4. Implementasi metode decision tree

Tahapan selanjutnya adalah implementasi metode *decision tree*. Tahapan ini bertujuan untuk menerapkan metode *decision tree* pada data set yang telah tersedia. Dalam impementasi metode *decision tree* peneliti menggunakan aplikasi WEKA 3.8.6 sebagai alat ujinya. Implementasi metode *decision tree* dengan menggunakan WEKA 3.8.6 seperti diperlihatkan oleh gambar 4.



Gambar 4 Implementasi Metode *Decision Tree* dengan WEKA 3.8.6

2.5. Evaluasi

Tahap yang terakhir adalah evaluasi dari hasil kinerja dari metode *decision tree*. Tahapan ini bertujuan untuk mengetahui kinerja dari algoritma *decision tree* pada data set yang tersedia. Berikut data hasil dari performa *decision tree* yang akan di evaluasi.

1. Clasifier output

Output dari klasifikasi data kualitas produk minuman A dengan algoritma *decision tree* seperti diperlihatkan oleh dambar 5.

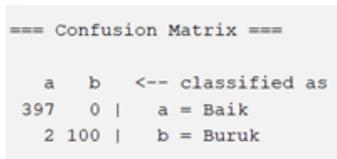
Classifier output								
Correctly Classified Instances	497 99.5992 %							
Incorrectly Classified Instances	2 0.4008 %							
Kappa statistic	0.9876							
Mean absolute error	0.0035							
Root mean squared error	0.0504							
Relative absolute error	1.0869 %							
Root relative squared error	12.4855 %							
Total Number of Instances	499							
=== Detailed Accuracy By Class ===								
TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
1.000	0.020	0.995	1.000	0.997	0.988	1.000	1.000	Baik
0.980	0.000	1.000	0.980	0.990	0.988	1.000	1.000	Buruk
Weighted Avg.	0.996	0.016	0.996	0.996	0.996	0.988	1.000	1.000

Gambar 5 Classifier Output

Gambar 5 diatas menunjukkan dengan jelas bahwa algoritma *decion tree* telah mampu melakukan klasifikasi dengan baik yakni dengan nilai akurasi sebesar 99.59 %. Ini berarti prediksi model untuk label data set sama dengan label sebenarnya yang ditetapkan untuk data set tersebut. Hasil klasifikasi juga memiliki nilai Kappa yang sangat baik yaitu sebesar 0.98. Nilai kappa dalam klasifikasi *decision tree* merupakan statistik yang digunakan untuk mengukur tingkat kecocokan antara prediksi model dan label sebenarnya dari data. Nilai ini dapat digunakan untuk mengevaluasi performa model dalam klasifikasi data, klasifikasi bernilai sempurna jika nilai Kappa nya mendekati 1.

2. *Confusion Matrix*

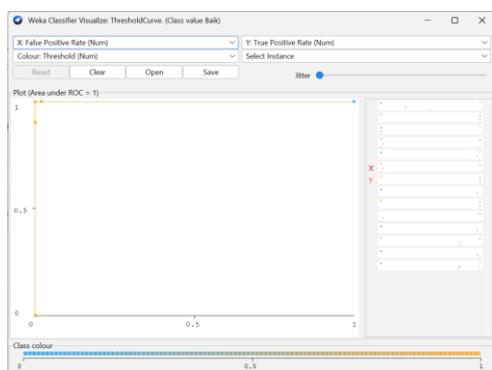
Evaluasi berikutnya adalah data dari confusion matrik dapat diketahui bahwa jumlah kelas status kualitas baik yang bernilai benar ada sebanyak 397 dan kelas status kualitas baik yang bernilai salah ada 2. Sedangkan jumlah untuk kelas status kualitas buruk yang bernilai benar ada 100 dan jumlah untuk kelas status kualitas buruk yang bernilai benar ada 0 , seperti yang ditunjukkan oleh gambar 6. Dari nilai *confusion matrix* tersebut akan dihasilkan nilai akurasi, presisi, recall dan *F1 score* yang bernilai baik seperti yang ditunjukkan oleh gambar 5 di atas.



Gambar 6 Confusion Matrix

3. ROC

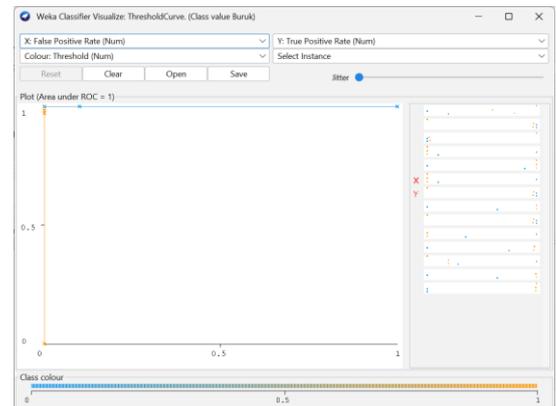
Kurva ROC pada penelitian ini menunjukkan bahwa model *decision tree* memiliki kemampuan yang baik dalam membedakan antara produk berkualitas baik dan produk yang berkualitas buruk. Berikut ini gambar dari kurva ROC hasil klasifikasi algoritma *decision tree*. Pada gambar 7 diketahui bahwa untuk kelas status kualitas baik bernilai sama dengan 1 hal ini berarti klasifikasi bernilai sempurna,



Gambar 7 Kurva ROC untuk Kelas Status Kualitas Baik

Untuk kelas status kualitas buruk menunjukkan nilai ROC sebesar 1 ini menunjukkan kelas Dari kurva ROC diatas menunjukkan bahwa model klasifikasi dapat

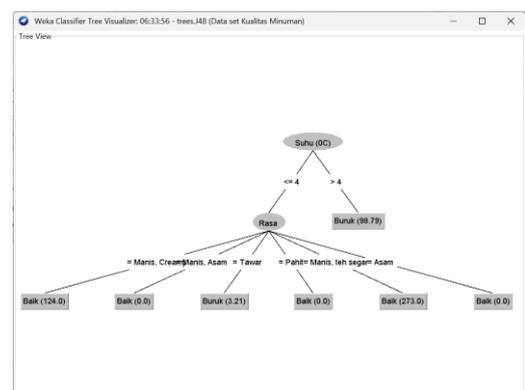
membedakan kelas positif dan negative dengan baik. Gravitik ROC untuk kelas kualitas buruk seperti diperlihatkan oleh gambar 8.



Gambar 8 Kurva ROC untuk Kelas Status Kualitas Buruk

4. *Decision Tree Visualization*

Visualisasi Pohon Keputusan adalah cara untuk merepresentasikan model pohon keputusan dalam format grafis. Pohon keputusan sendiri merupakan model pembelajaran mesin yang menggunakan struktur mirip pohon untuk mengklasifikasikan data atau memprediksi nilai target. Mereka memulai dengan simpul akar yang berisi seluruh kumpulan data. Setiap node internal mengajukan pertanyaan tentang atribut (fitur) tertentu dari data. Berdasarkan jawabannya (benar atau salah), data dibagi menjadi dua cabang. Proses ini berlanjut hingga mencapai *node* daun, yang mewakili prediksi akhir (klasifikasi atau nilai) untuk suatu titik data. Dari gambar 9 dibawah ini visual *tree* yang terbentuk pada penelitian ini sangat jelas jika atribut yang paling berperan dalam melakukan klasifikasi kualitas produk minuman A adalah atribut suhu, selanjutnya diikuti oleh atribut rasa.



Gambar 9 Visualisasi Tree Algoritma *Decision Tree*

3. Kesimpulan

Penelitian ini menunjukkan bahwa metode *decision tree* dapat diimplementasikan dengan sukses pada sistem prediksi status kualitas produk minuman A di UMKM BUAH SABAR. Model *decision tree* yang dibangun mampu memprediksi status kualitas produk dengan

tingkat akurasi yang tinggi, yaitu 99,59%. Atribut yang paling penting dalam menentukan status kualitas produk adalah suhu dan rasa, atribut-atribut ini memiliki pengaruh yang signifikan terhadap kualitas produk dan dapat digunakan sebagai parameter utama dalam sistem prediksi. Sistem prediksi ini memiliki beberapa manfaat, antara lain meningkatkan efisiensi dan efektivitas proses kontrol kualitas produk, membantu UMKM BUAH SABAR dalam mengambil keputusan yang tepat terkait dengan kualitas produk, meningkatkan daya saing produk di pasaran. Penelitian ini menunjukkan bahwa metode *decision tree* merupakan alat yang bermanfaat dan efektif untuk memprediksi status kualitas produk minuman. Penerapan metode ini dapat membantu UMKM BUAH SABAR dalam meningkatkan kualitas dan daya saing produknya.

Daftar Pustaka

- Afdhal, Muhammad, Vicky Ariandi, and Rita Rita. 2022. "Memprediksi Penjualan Pada Toko Hanifah Metode C.45." *Jurnal Teknologi Dan Sistem Informasi Bisnis* 4(2):248–55. doi: 10.47233/jteksis.v4i1.460.
- Firdausia Ismi Nurhayati, Budi Nugroho, and Intan Yuniar Purbasari. 2021. "Implementasi Metode Decision Tree Pada Identifikasi Status Gizi Balita." *Jurnal Informatika Dan Sistem Informasi* 2(2):204–13. doi: 10.33005/jifosi.v2i2.326.
- Hts, Dedek Indra Gunawan. 2018. "Penerapan Algoritma C4.5 Untuk Penerimaan Siswa Baru Pada SMA Al-Azhar Medan." *J-SAKTI (Jurnal Sains Komputer Dan Informatika)* 2(2):130. doi: 10.30645/j-sakti.v2i2.76.
- Irnawati, Oky, and Nia Nuraeni. 2022. "Optimasi Decision Tree Dengan Particle Swarm Optimization Untuk Klasifikasi Prediksi Gender Pengguna E-Commerce." *Journal of Information Engineering and Educational Technology* 6(1):37–41. doi: 10.26740/jieet.v6n1.p37-41.
- Lutfi, Moch. 2018. "KLASIFIKASI KUALITAS MUTU JAGUNG DENGAN MENGGUNAKAN METODE DECISION TREE PADA DINAS PERTANIAN BOJONEGORO." *Agromix* 9(1). doi: 10.35891/agx.v9i1.1360.
- Negara, Yudha Dwi Putra, Fifin Ayu Murfarroha, Vina Angelina Savitri, Ach. Khozaimi, and Budi Soesilo. 2023. "PENERAPAN METODE DECISION TREE UNTUK KLASIFIKASI DATA PRODUK SKINCARE UNTUK IBU HAMIL MENGGUNAKAN APLIKASI ORANGE." *Insand Comtech: Information Science and Computer Technology Journal* 7(2). doi: 10.53712/jic.v7i2.1680.
- Rahman, Fauzi Yusa, Indu Indah Purnomo, and Nadya Hijriana. 2022. "PENERAPAN ALGORITMA DATA MINING UNTUK KLASIFIKASI KUALITAS AIR." *Technologia : Jurnal Ilmiah* 13(3):228. doi: 10.31602/tji.v13i3.7070.
- Rewo Ngeo, Sandrianus and Latipah. 2021. "PENERAPAN METODE DECISION TREE UNTUK REKOMENDASI TUJUAN POLI PADA RUMAH SAKIT UMUM DAERAH BAJAWA." *Antivirus : Jurnal Ilmiah Teknik Informatika* 15(1):62–74. doi: 10.35457/antivirus.v15i1.1289.
- Sahputra, Deni Anugrah, M. Rangga Ramadhan Saellan, Lilyani Asri Utami, and Windu Gata. 2020. "Penentuan Faktor Kelayakan Penerimaan Karyawan Menggunakan Algoritma Decision Tree Pada Perusahaan PT. Personel Alih Daya." *Jurnal Sains Dan Informatika* 6(2):148–56. doi: 10.34128/jsi.v6i2.225.
- Setiawati, Intan, Adityo Permana, and Arief Hermawan. 2019. "IMPLEMENTASI DECISION TREE UNTUK MENDIAGNOSIS PENYAKIT LIVER." *Journal of Information System Management (JOISM)* 1(1):13–17. doi: 10.24076/JOISM.2019v1i1.17.
- Wicaksono, Arif Wicaksono, and Tedy Setiadi. 2023. "Penerapan Klasifikasi Decision Tree (C4.5) Untuk Memprediksi Kelulusan Siswa Sekolah Dasar Di Kecamatan Juai." *Format : Jurnal Ilmiah Teknik Informatika* 12(2):151. doi: 10.22441/format.2023.v12.i2.008.