

Penerapan Finite State Automata Pada Vending Machine Parfum Laundry Pakaian Dan Karpet

Yessica Fara Desvia^{1*}, Febryawan Yuda Pratama¹, Suhendri¹

¹Bisnis Digital, Politeknik Jatiluhur, Purwakarta, Indonesia

E-mail: yessicadesvia@polijati.ac.id^{1*}, febryawanpratama@polijati.ac.id, suhendri@polijati.ac.id

Received: 2025-07-11 | Accepted: 2025-09-02 | Published: 2025-10-01

Abstrak

Parfum digemari oleh berbagai lapisan masyarakat, salah satunya adalah jenis parfum yang digunakan dalam proses *laundry*. Parfum *laundry* tersedia dalam beragam varian aroma, seperti aroma buah-buahan, bunga, kombinasi buah dan bunga, serta aroma kayu. Parfum ini biasanya digunakan pada tahap akhir proses pencucian. Saat ini, pelanggan menerima hasil *laundry* dengan aroma parfum yang ditentukan secara acak berdasarkan ketersediaan dari pihak penyedia jasa, sehingga mereka tidak dapat memilih aroma sesuai keinginan. Oleh karena itu, dibutuhkan rancangan sebuah *Vending Machine* (VM) yang memungkinkan pelanggan memilih parfum secara mandiri. Rancangan VM ini menggunakan pendekatan *Finite State Automata* (FSA) dengan tipe *Non-Deterministic Finite Automata* (NFA), karena metode ini mampu menangani beberapa kondisi dalam satu pilihan. Pengembangan metode NFA melibatkan tahapan analisis proses bisnis, pembuatan diagram state, perancangan VM, serta pengujian sistem. Hasil penelitian menunjukkan bahwa kehadiran VM ini memudahkan pelanggan dalam memilih aroma parfum untuk *laundry* mereka, sehingga hasil cucian memiliki aroma yang sesuai dengan preferensi masing-masing pelanggan.

Kata Kunci: *Finite State Automata, Parfum Laundry, Vending Machine.*

Abstract

Perfume is popular among various groups of people, including those who use laundry fragrances. Laundry perfumes come in a variety of scents, such as fruity, floral, a combination of fruit and floral, and woody aromas. These fragrances are typically applied during the final stage of the laundry process. Currently, customers receive their laundry with a randomly selected scent based on the availability at the laundry service, which means they cannot choose the fragrance they prefer. Therefore, a *Vending Machine* (VM) design is needed to allow customers to select their desired laundry perfume. The VM is designed using the *Finite State Automata* (FSA) approach, specifically the *Non-Deterministic Finite Automata* (NFA) type, as it can accommodate multiple conditions for a single option. The development of the NFA method involves several stages, including business process analysis, state diagram creation, VM design, and system testing. The results of this study indicate that implementing this VM simplifies the process for customers to choose their preferred laundry perfume, ensuring that their laundry has a scent that matches their personal preferences.

Keywords: *Finite State Automata, Parfum Laundry, Vending Machine.*

1. Pendahuluan

Dalam falsafah Jawa, kebutuhan utama yang pertama kali disebut adalah sandang atau pakaian, kemudian diikuti oleh kebutuhan lainnya. Pakaian yang bersih, rapi, dan harum sangat mendukung penampilan seseorang, sehingga masyarakat memberikan perhatian khusus terhadap pakaian yang dikenakan. Untuk memberikan kenyamanan saat berinteraksi sosial, penggunaan parfum menjadi hal yang penting.

Parfum sendiri merupakan produk yang sangat familiar dalam kehidupan sehari-hari. Setiap individu memiliki preferensi aroma yang berbeda—mulai dari aroma bunga, buah-buahan, perpaduan keduanya, hingga aroma kayu dan lainnya. Untuk menciptakan aroma yang tahan lama, digunakan bahan tambahan yang disebut fixative sebagai pengikat parfum. Hal ini penting karena konsentrasi parfum murni memiliki tingkat volatilitas tinggi, yang menyebabkan wanginya cepat menguap dari pakaian. Dengan adanya fixative, volatilitas tersebut dapat dikurangi, sehingga aroma dapat bertahan lebih lama[1],[2].

Salah satu jenis parfum yang digunakan adalah parfum laundry, yang diaplikasikan pada tahap akhir proses pencucian pakaian. Layanan laundry sangat diminati masyarakat. Berdasarkan survei dari media cetak, segmen terbesar pengguna laundry kiloan adalah mahasiswa (48%), diikuti oleh pekerja (32%), ibu rumah tangga (16%), dan lainnya (4%)[3].

Bagian laundry atau laundry section merupakan unit yang bertanggung jawab atas pencucian seluruh jenis pakaian baik milik tamu maupun internal (seperti linen hotel). Saat ini, laundry juga melayani pelanggan eksternal untuk meningkatkan keuntungan[4]

Bagi masyarakat dengan rutinitas padat, layanan laundry menjadi alternatif praktis karena seringkali disertai dengan layanan setrika sekaligus, sehingga membantu meringankan beban mereka yang tidak sempat mencuci atau menyetrika pakaian sehari-hari.

Layanan laundry kini tersedia baik secara langsung (manual) maupun daring, dan umumnya berjalan dengan baik. Namun, salah satu kelemahan yang masih ada adalah pelanggan belum bisa memilih aroma parfum sesuai keinginan. Umumnya, mereka hanya menerima hasil cucian yang bersih dan harum, tetapi dengan aroma yang dipilih oleh pihak laundry, sehingga kepuasan mereka belum optimal.

Untuk menunjang perkembangan industri laundry, salah satu inovasi yang bisa diterapkan adalah penggunaan Vending Machine (VM). Dengan adanya VM, pelanggan dapat memilih sendiri aroma parfum laundry sesuai keinginan mereka. Umumnya, VM berbentuk kotak logam dengan tampilan depan transparan agar produk di dalamnya terlihat. Selama ini, VM banyak digunakan untuk menjual makanan dan minuman kemasan, tiket, dan lain sebagainya[5]

Penelitian-penelitian sebelumnya menunjukkan bahwa VM telah dirancang untuk berbagai kebutuhan, seperti penjualan susu kambing etawa dengan berbagai rasa menggunakan metode Finite State Automata (FSA) jenis Non-Deterministic Finite Automata (NFA). VM ini terbukti membantu distribusi produk terutama di kota besar[5].

VM juga pernah diterapkan untuk penjualan rujak buah menggunakan metode FSA jenis NFA. Desain VM tersebut memungkinkan pelanggan memilih berbagai jenis buah dan tingkat kepedasan sesuai selera, sehingga berkontribusi pada peningkatan nilai jual dan daya tarik produk[6].

Penelitian lain menunjukkan penerapan VM dalam penjualan jamu dan yoghurt, juga dengan metode FSA jenis NFA. Hasilnya menunjukkan bahwa metode FSA dapat menjadi dasar yang kuat untuk membangun sistem VM dengan desain efisien, seperti penggunaan jumlah state yang minimal dan dukungan transaksi dengan uang kertas. [7], [8], [9]

Studi lain juga merancang VM untuk penjualan kopi dengan metode NFA, di mana proses peracikan dan pembayaran dilakukan secara otomatis untuk menghasilkan produk yang sesuai permintaan pelanggan[10]. Tidak hanya terbatas pada makanan dan minuman, VM juga dirancang untuk penjualan rokok yang terintegrasi dengan e-KTP sebagai upaya membatasi akses rokok bagi anak di bawah umur[11].

Penelitian ini fokus pada perancangan VM untuk pemilihan parfum *laundry*—topik yang belum pernah dibahas sebelumnya. VM yang dirancang dalam penelitian ini memungkinkan pelanggan memilih parfum *laundry* saat memesan layanan *laundry*, dengan sistem berbasis metode FSA jenis NFA. Selain memungkinkan pemilihan aroma, sistem ini juga mendukung metode pembayaran tunai maupun non-tunai.

Tujuan dari penelitian ini adalah untuk menyelesaikan permasalahan pelanggan *laundry* yang tidak dapat memilih aroma parfum sesuai preferensi mereka. Dengan adanya VM ini, pelanggan dapat menentukan sendiri wangi yang mereka inginkan, yang pada akhirnya meningkatkan kepuasan. Penelitian ini menghasilkan desain VM pemilihan parfum *laundry* berbasis metode FSA jenis NFA, yang diharapkan dapat menjadi solusi praktis dan mendukung perkembangan industri *laundry* ke depannya.

2. Metode Penelitian

2.1. Metodologi Penelitian

1) Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif dan eksperimental. Penelitian ini bertujuan untuk menganalisis kelayakan penerima Program Keluarga Harapan (PKH) menggunakan dua algoritma data mining, yaitu Naïve Bayes dan C4.5.

2) Sumber Data

Data yang digunakan dalam penelitian ini diperoleh dari survei lapangan di desa yang mencakup informasi sosial dan ekonomi keluarga penerima PKH, termasuk penghasilan, jumlah anggota keluarga, status pekerjaan, pendidikan, dan kondisi kesehatan. Data yang telah terkumpul kemudian divalidasi dan dikonsolidasikan menjadi dataset yang dapat digunakan untuk analisis.

3) Teknik Pengolahan Data

Data yang telah diperoleh melalui survei akan diproses dan dianalisis menggunakan algoritma Naïve Bayes dan C4.5 untuk mengklasifikasikan keluarga yang layak menerima bantuan berdasarkan kriteria yang telah ditentukan.

- Naïve Bayes digunakan untuk mengklasifikasikan keluarga berdasarkan probabilitas kondisi sosial dan ekonomi.
- C4.5 digunakan untuk membangun pohon keputusan yang dapat menggambarkan hubungan antara variabel-variabel yang mempengaruhi kelayakan penerima bantuan.

4) Langkah-langkah Analisis

- Persiapan Data: Data yang diperoleh melalui survei lapangan akan disaring dan disusun untuk memenuhi format yang diperlukan oleh algoritma.
- Pengolahan Data: Setelah data dipersiapkan, algoritma Naïve Bayes dan C4.5 akan diterapkan untuk menganalisis kelayakan penerima PKH.
- Evaluasi Model: Hasil dari kedua algoritma akan dievaluasi berdasarkan tingkat akurasi dan efisiensi dalam mengklasifikasikan penerima bantuan.

5) Evaluasi dan Pengujian

Model yang dibangun akan diuji untuk menilai akurasi prediksi penerima bantuan dengan membandingkan hasil klasifikasi dengan data sebenarnya (ground truth). Evaluasi dilakukan dengan menggunakan akurasi, presisi, recall, dan F1-score.

2.2. Tahapan Pemodelan Menggunakan Algoritma Naïve Bayes dan C4.5:

a. Teknik Pengambilan Data

Pengambilan data untuk Program Keluarga Harapan (PKH) di desa melibatkan beberapa langkah untuk memastikan data yang akurat, yaitu:

- Pendataan Awal: Melakukan survei lapangan, wawancara dengan kepala keluarga atau anggota keluarga, dan observasi kondisi rumah serta aset keluarga untuk memperoleh informasi sosial dan ekonomi.
- Validasi Data: Data yang terkumpul divalidasi untuk memastikan akurasi, melalui pengecekan ulang atau menggunakan data dari instansi terkait.
- Konsolidasi Data: Data yang telah diverifikasi kemudian disusun menjadi dataset penerima PKH yang valid.

b. Penentuan Kriteria Penerima

Berdasarkan data yang dikumpulkan, kriteria kelayakan penerima PKH ditentukan, seperti jumlah anggota keluarga, penghasilan, pendidikan, status rumah, pekerjaan, dan kondisi kesehatan. Algoritma Naïve Bayes atau C4.5 kemudian digunakan untuk mengklasifikasikan keluarga yang layak menerima bantuan, berdasarkan kriteria tersebut.

2.3. Pemodelan Sistem

a. Use Case Diagram

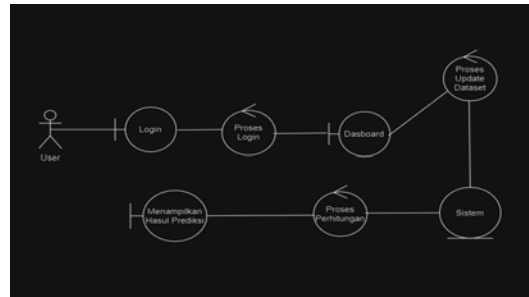
Berikut diagram *use case* gambar 3.1 pada aplikasi Penerapan Algoritma Naïve Bayes dan C4.5 pada Sistem Perbandingan klasifikasi data PKH.



Gambar 3.1 Use Case Diagram

b. *Communication Diagram*

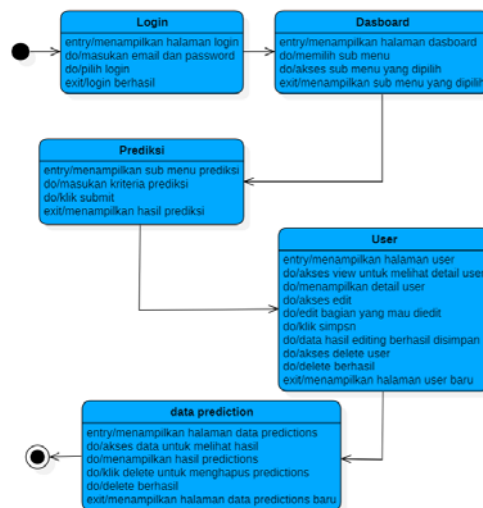
Berikut diagram *communication* gambar 3.2 pada aplikasi Penerapan Algoritma Naïve Bayes dan C4.5 pada Sistem Perbandingan klasifikasi data PKH (studi kasus: Desa Tenjolaya).



Gambar 3.2 *Communication Diagram*

c. *Statechart Diagram*

Berikut diagram *statechart* gambar 3.3 pada aplikasi Penerapan Algoritma Naïve Bayes dan C4.5 pada Sistem Perbandingan klasifikasi data PKH (studi kasus: Desa Tenjolaya).



Gambar 3.3 *Statechart Diagram*

2.4. Perhitungan Algoritma Naive Bayes

Berikut ini adalah perhitungan manual penyelesaian menggunakan algoritma Naïve Bayes dalam Klasifikasi kelayakan.

1. Menghitung Probabilitas Prior

setiap kategori, yang menjadi kategori ada 6 kategori yaitu kelas jumlah keluarga, penghasilan per bulan, pendidikan tertinggi, status rumah, pekerjaan, kondisi Kesehatan

Untuk menghitung probabilitas Prior dengan rumus

$$Prob = \frac{frekuensi}{jumlah. data}$$

Probabilitas prior untuk masing-masing anggota yaitu:

$$P(\text{status penerima} = \text{ya}) = 0,50$$

$$P(\text{status penerima} = \text{tidak}) = 0,50$$

2. Menghitung Probabilitas Kondisi (Likelihood)

melibatkan enam variabel yaitu jumlah keluarga, penghasilan per bulan, pendidikan tertinggi, status rumah, pekerjaan, dan kondisi kesehatan, di mana perhitungan probabilitas kondisi sama seperti probabilitas prior tetapi melibatkan anggota atau kelas target.

Rumus Probabilitas Kondisi (Likelihood)

Probabilitas kondisi untuk setiap kategori dihitung dengan membagi frekuensi kategori tersebut dengan total frekuensi dalam kelas yang sama.

Probabilitas "Ya" untuk setiap kategori:

Probabilitas "Ya" untuk "kecil":

$$Probabilitas (kecil | Ya) = \frac{\text{Frekuensi kecil (Ya)}}{\text{Total Frekuensi (Ya)}} = \frac{5}{30} = 0,16$$

Probabilitas "Ya" untuk "sedang":

$$Probabilitas (sedang | Ya) = \frac{\text{Frekuensi sedang (Ya)}}{\text{Total Frekuensi sedang (Ya)}} = \frac{24}{30} = 0,8$$

Probabilitas "Ya" untuk "besar":

$$Probabilitas (besar | Ya) = \frac{\text{Frekuensi besar (Ya)}}{\text{Total Frekuensi sedang (Ya)}} = \frac{1}{30} = 0,03$$

Probabilitas "Tidak" untuk setiap kategori:

Probabilitas "Tidak" untuk "kecil":

$$Probabilitas (kecil | Tidak) = \frac{\text{Frekuensi kecil (tidak)}}{\text{Total Frekuensi (tidak)}} = \frac{10}{30} = 0,33$$

Probabilitas "Tidak" untuk "Sedang":

$$Probabilitas (sedang | Tidak) = \frac{\text{Frekuensi sedang (tidak)}}{\text{Total Frekuensi (tidak)}} = \frac{10}{30} = 0,33$$

Probabilitas "Tidak" untuk "besar":

$$Probabilitas (besar | Tidak) = \frac{\text{Frekuensi besar (tidak)}}{\text{Total Frekuensi (tidak)}} = \frac{10}{30} = 0,33$$

1. Probabilitas kondisi jumlah keluarga

Tabel 3.1 Probabilitas Jumlah Keluarga

Anggota jumlah keluarga	Frekuensi		Probabilitas	
	Ya	Tidak	Ya	Tidak
kecil	5	10	0,16	0,33
Sedang	24	10	0,8	0,33
besar	1	10	0,03	0,33
Total	30	30	1	1

2. Probabilitas kondisi pengasilan

Tabel 3.2 Probabilitas Penghasilan

penghasilan per bulan	Frekuensi		Probabilitas	
	Ya	Tidak	Ya	Tidak
kurang	23	0	0,76	0,0
Sedang	9	9	0,3	0,3
tinggi	0	23	0,0	0,76
Total	30	30	1	1

penghasilan per bulan kecil

$$P(\text{penghasilanPerBulan=kurang} = ya) = 0,69$$

$$P(\text{penghasilanPerBulan=kurang} = tidak) = 0,03$$

penghasilan per bulan sedang

$$P(\text{penghasilanPerBulan=sedang} = ya) = 0,27$$

$$P(\text{penghasilanPerBulan=sedang} = tidak) = 0,27$$

penghasilan per bulan tinggi

$$P(\text{penghasilanPerBulan=tinggi} = ya) = 0,03$$

$$P(\text{penghasilanPerBulan=tinggi} = tidak) = 0,69$$

3. Probabilitas kondisi Pendidikan

Tabel 3.3 Probabilitas Pendidikan

pendidikan tertinggi	Frekuensi		Probabilitas	
	Ya	Tidak	Ya	Tidak
sd	15	6	0,42	0,17
smp	13	13	0,37	0,37
sma	1	9	0,02	0,25
tidak sekolah	5	4	0,14	0,11
sarjana	1	3	0,02	0,08
Total	35	35	1	1

4. Probabilitas setatus rumah

Tabel 3.4 Probabilitas Pendidikan

setatus rumah	Frekuensi		Probabilitas	
	Ya	Tidak	Ya	Tidak
kurang	8	5	0,24	0,15
Sedang	24	13	0,72	0,39
baik	1	15	0,03	0,45
Total	33	33	1	1

5. Probabilitas setatus pekerjaan

Tabel 3.5 Probabilitas Pekerjaan

pekerjaan	Frekuensi		Probabilitas	
	Ya	Tidak	Ya	Tidak
lain-lain	23	20	0,69	0,60
petani	9	6	0,27	0,18
wirusaha	1	7	0,03	0,21
Total	33	33	1	1

6. Probabilitas setatus Kesehatan

Tabel 3.6 Probabilitas kesehatan

setatus kesehatan	Frekuensi		Probabilitas	
	Ya	Tidak	Ya	Tidak
kurang	8	4	0,24	0,12
cukup	24	15	0,72	0,45
baik	1	14	0,03	0,42
Total	33	33	1	1

3. Perhitungan Klasifikasi

Setelah didapatkan nilai untuk probabilitas prior dan kondisi maka untuk klasifikasi pada Perhitungan Manual Naives Bayes sebagai berikut.

Untuk menghitung probabilitas suatu kategori (Ya/Tidak) berdasarkan kondisi-kondisi yang diberikan, kita menggunakan rumus sebagai berikut:

$$P(Ya | Kondisi) = \frac{P(Kondisi | Ya) \times P(Ya)}{P(Kondisi)}$$

$$P(Tidak | Kondisi) = \frac{P(Kondisi | Tidak) \times P(Tidak)}{P(Kondisi)}$$

Namun, karena $P(Kondisi)$ sama untuk kedua kasus, kita dapat membandingkan $P(Kondisi | Ya) \times P(Ya)$ dengan $P(Kondisi | Tidak) \times P(Tidak)$.

3.1 Tingkat Akurasi

Untuk menghitung tingkat Accuracy, Precision, Recall, dan F1-Score, kita perlu terlebih dahulu menghitung jumlah True Positives (TP), True Negatives (TN), False Positives (FP), dan False Negatives (FN) berdasarkan data.

Menentukan TP, TN, FP, dan FN

- True Positives (TP): Kategori yang benar dan diprediksi benar.
- True Negatives (TN): Kategori yang benar dan diprediksi benar.
- False Positives (FP): Kategori yang salah namun diprediksi benar.
- False Negatives (FN): Kategori yang benar namun diprediksi salah.

Tabel 3.7 Tingkat Akurasi

no	ya	tidak	fakta	klasifikasi	prediksi	Kategori yang	Kategori yg	True/False
						Benar	Diprediksi	
1	0.00119113	0.00046286	Ya	Ya	sesuai	Ya	Ya	TP
2	0.00000037	0.00015324	Tidak	Tidak	sesuai	Tidak	Tidak	TN

no	ya	tidak	fakta	klasifikasi	prediksi	Kategori yang	Kategori yg	True/False
						Benar	Diprediksi	Positive/Negative
3	0.01217597	0.00006743	Ya	Ya	sesuai	Ya	Ya	TP
4	0.01382137	0.00005212	Ya	Ya	sesuai	Ya	Ya	TP
5	0.00000000	0.00102698	Tidak	Tidak	sesuai	Tidak	Tidak	TN
....								
....								
59	0.00000000	0.00110003	Tidak	Tidak	sesuai	Tidak	Tidak	TN
60	0.00000000	0.00293423	Tidak	Tidak	sesuai	Tidak	Tidak	TN

Jumlah TP, TN, FP, FN:

- TP (True Positives): 26
- TN (True Negatives): 26
- FP (False Positives): 0
- FN (False Negatives): 0

1. Menghitung Numerik

1. Accuracy:

$$\text{Accuracy} = \frac{TP+TN}{\text{Total}}$$

$$\text{Accuracy} = \frac{26+26}{60} = \frac{52}{60} = 0.867$$

2. Precision:

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Precision} = \frac{26}{26+0} = 1$$

3. Recall:

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{Recall} = \frac{26}{26+0} = 1$$

4. F1-Score:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{F1-Score} = 2 \times \frac{1 \times 1}{1+1} = 1$$

Hasil Metrik:

- Accuracy: 0.980 (98%)
- Precision: 1 (100%)
- Recall: 1 (100%)
- F1-Score: 1 (100%)

Data menunjukkan kinerja yang sangat baik dengan precision, recall, dan F1-Score masing-masing 100%, dan akurasi 98%.

3.8 Perhitungan Algoritma C4.5

NO	Jumlah Keluarga	Penghasilan	Pendidikan	Sts Rumah	Pekerjaan	Kesehatan	Sts Penerima
1	Keluarga Kecil	Sedang	SMP	Sedang	lain-lain	Kurang	Ya
2	Keluarga Besar	Sedang	Tidak Sekolah	Sedang	Wirausaha	Baik	Tidak
3	Keluarga Sedang	Kurang	SMP	Kurang	lain-lain	Cukup	Ya
4	Keluarga Sedang	Kurang	Tidak Sekolah	Sedang	lain-lain	Cukup	Ya
5	Keluarga Kecil	Tinggi	SMA	Baik	Wirausaha	Baik	Tidak
....							

NO	Jumlah Keluarga	Penghasilan	Pendidikan	Sts Rumah	Pekerjaan	Kesehatan	Sts Penerima
59	Keluarga Kecil	Tinggi	SMA	Baik	Wirausaha	Baik	Tidak
60	Keluarga Sedang	Tinggi	SMP	Baik	lain-lain	Baik	Tidak

Mengelompokkan setiap variabel untuk mendapatkan atribut yang akan digunakan untuk menghitung entropy dari kasus agar mendapatkan akar dan cabang dari penerima bantuan PKH.

Tabel 3.9 Perhitungan manual

keterangan	Jumlah Kasus	Status (Ya)	Status (Tidak)	Entropy	Gain
Total	60	30	30	1	
Jumlah Keluarga					0,1945957
Keluarga Kecil	15	5	10	0,918295834	
Keluarga Sedang	34	24	10	0,873981048	
Keluarga Besar	11	1	10	0,439496987	
Penghasilan					0,7509775
Kurang	22	23	0	0	
Sedang	16	9	9	0,933834373	
Tinggi	22	0	23	0	
Pendidikan					0,2217565
SD	19	14	5	0,831474388	
SMP	24	12	12	1	
SMA	8	0	8	0	
Tidak Sekolah	7	4	3	0,985228136	
SARJANA	2	0	2	0	
Status Rumah					0,2855718
Kurang	11	7	4	0,945660305	
Sedang	35	23	12	0,927526588	
Baik	14	0	14	0	
Pekerjaan					0,3527269
lain-lain	41	22	19	0,996134484	
petani	6	8	5	-0,33418799	
Wirausaha	13	0	6	0	
Kesehatan					0,2630368
Kurang	10	7	3	0,881290899	
Cukup	37	23	14	0,956888666	
Baik	13	0	13	0	

1. menghitung nilai entropy dari semua atribut

Berikut adalah nilai entropy:

Pertama adalah mencari entropy dari total data yang ada, jumlah data yang digunakan adalah 60 data terbagi menjadi 30 “Ya” dan 30 “Tidak”, berikut contoh perhitungannya:

a. Entropy Total

$$= Entropy(S) \sum_{i=1}^n (-P_i \log_2 P_i)$$

$$= ((-30/60) * \log_2(30/60) + (-30/60) * \log_2(30/60))$$

$$= 1$$

b. Entropy Jumlah Keluarga

- Keluarga kecil = 0,918295834
- keluarga Sedang = 0,873981048
- Keluarga Besar = 0,439496987

c. Entropy Penghasilan

- Kurang = 0
- Sedang = 0,933834373

- Tinggi = 0
 - d. *Entropy* Pendidikan
 - Sd = 0,831474388
 - Smp = 1
 - Sma = 0
 - Tidak Sekolah = 0,985228136
 - Sarjana = 0
 - e. *Entropy* Setatus Rumah
 - Kurang = 0,945660305
 - Sedang = 0,927526588
 - Baik = 0
 - f. *Entropy* Pekerjaan
 - Lain-lain = 0,996134484
 - Petani = 0,961236605
 - Wirausaha = 0
 - g. *Entropy* Kesehatan
 - Kurang = 0,881290899
 - Cukup = 0,956888666
 - Baik = 0
2. Menghitung Nilai Gain Dari Semua Atribut
Berikut adalah nilai gain dari setiap atribut:

$$Gain = Entropy(S) - \sum_{i=1}^{|S|} \frac{|S_i|}{|S|} * Entropy(S_i)$$

- a. Nilai Gain Jumlah Keluarga
= 0,194595667
- b. Nilai Gain Penghasilan
= 0,7509775
- c. Nilai Gain Pendidikan
= 0,221756495
- d. Nilai Gain Setatus Rumah
= 0,285571768
- e. Nilai Gain Pekerjaan
= 0,111040172
- f. Nilai Gain Kesehatan
= 0,26303684

Berdasarkan perhitungan diatas dipilih atribut yang memiliki nilai gain tertinggi untuk dijadikan sebagai akar. Pada tabel 3.72 terlihat atribut penghasilan akan dipilih sebagai akar. Terdapat 3 atribut pada penghasilan yaitu kurang, sedang, tinggi. dari ketiga atribut tersebut, nilai atribut kurang sudah mengklasifikasikan kasus menjadi 1 yaitu keputusannya Ya dan atribut kurang sudah mengklasifikasikan kasus menjadi Tidak, sehingga tidak perlu dilakukan perhitungan lagi. Untuk nilai atribut sedang masih perlu dilakukan perhitungan lagi.

Selanjutnya menghitung nilai dari atribut jumlah keluarga yang memiliki nilai gain lebih besar dan hasil perhitungan dapat dilihat pada tabel

Tabel 3.10 Node 1.1

keterangan	Jumlah Kasus	Status (Ya)	Status (Tidak)	Entropy	Gain
Total	60	30	30	1	
Jumlah Keluarga					0,8622556
Keluarga Kecil	2	2	0	0	
Keluarga Sedang	9	6	3	0,918295834	
Keluarga Besar	5	0	5	0	
Pendidikan					0,7883856
SD	6	5	1	0,650022422	
SMP	7	2	5	0,863120569	
SMA	0	0	0	0	

keterangan	Jumlah Kasus	Status (Ya)	Status (Tidak)	Entropy	Gain
Tidak Sekolah	3	1	2	0,918295834	
SARJANA	0	0	0	0	
Status Rumah					0,7917321
Kurang	1	0	1	0	
Sedang	13	8	5	0,961236605	
Baik	2	0	2	0	
Pekerjaan					0,7513386
lain-lain	9	5	4	0,99107606	
petani	6	3	3	1	
Wirausaha	1	0	1	0	
Kesehatan					0,7748371
Kurang	6	4	2	0,918295834	
Cukup	8	4	4	1	
Baik	2	0	2	0	

Selanjutnya menghitung nilai dari atribut pendidikan untuk dijadikan node percabangan dari nilai atribut sd, smp dan tidak sekolah. Perhitungan dilakukan dengan cara sama dan hasil perhitungan dapat dilihat pada tabel 3.11.

Tabel 3.11 Node 1.111

keterangan	Jumlah Kasus	Status (Ya)	Status (Tidak)	Entropy	Gain
Total	60	30	30	1	
Pendidikan					0,4034203
SD	18	14	4	0,764204507	
SMP	18	12	6	0,918295834	
SMA	3	0	3	0	
Tidak Sekolah	6	4	2	0,918295834	
SARJANA	1	0	1	0	
Pekerjaan					0,3244838
lain-lain	35	22	13	0,951762676	
petani	10	8	2	0,721928095	
Wirausaha	1	0	1	0	
Kesehatan					0,3959724
Kurang	10	7	3	0,881290899	
Cukup	32	23	9	0,857148437	
Baik	4	0	4	0	

Dari hasil perhitungan entropy dan information gain yang didapat kemudian diolah kedalam decision tree, berikut hasil pohon keputusan (Decision Tree).

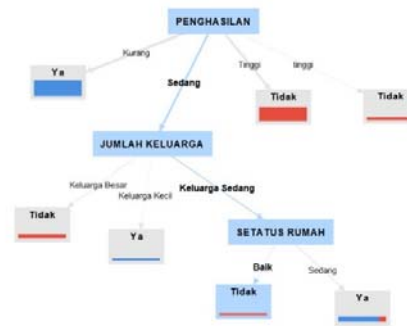
1. Entropy dan Gain:

Tabel 3.12 Entropy dan Gain

Atribut	Entropy	Gain
Jumlah Keluarga	0,918 - 0,439	0,1946
Penghasilan	0 - 0,933	0,751
Pendidikan	0 - 0,985	0,2218
Status Rumah	0 - 0,946	0,2856
Pekerjaan	0 - 0,996	0,3527
Kesehatan	0 - 0,956	0,263

2. Pohon Keputusan C4.5:

Berdasarkan nilai gain tertinggi, atribut Penghasilan dengan gain 0,7509775 menjadi akar (root) dari pohon keputusan, diikuti oleh atribut lain yang memiliki gain tertinggi berikutnya.



3.4 Pohon Keputusan

3. Evaluasi Model:

Jika model sudah dibangun dan diuji dengan data uji (testing data), kita bisa menghitung Accuracy, Precision, Recall, dan F1-Score.

4. Menghitung Accuracy, Precision, Recall, dan F1-Score:

Untuk menghitung metrik tersebut, Anda memerlukan informasi mengenai:

- True Positives (TP): Kasus yang diprediksi "Ya" dan memang benar "Ya".
- True Negatives (TN): Kasus yang diprediksi "Tidak" dan memang benar "Tidak".
- False Positives (FP): Kasus yang diprediksi "Ya" tetapi seharusnya "Tidak".
- False Negatives (FN): Kasus yang diprediksi "Tidak" tetapi seharusnya "Ya".

$$\text{Accuracy} \left(\frac{25 + 28}{25 + 28 + 5 + 2} \right) = \left(\frac{53}{60} \right) = 93\%$$

$$\text{precision} \left(\frac{25}{25 + 5} \right) = \left(\frac{25}{30} \right) = 83,3\%$$

$$\text{Recall} \left(\frac{25}{25 + 2} \right) = \left(\frac{25}{27} \right) = 92,6\%$$

$$\text{F1 - Score} \left(\frac{2 \times 0,833 \times 0,926}{0,833 + 0,926} \right) = 87,7\%$$

5. Support:

Support adalah jumlah kasus aktual dari setiap kelas dalam dataset. Misalnya, dalam data Anda, terdapat 30 kasus "Ya" dan 30 kasus "Tidak".

Ringkasan:

Hasil evaluasi model C4.5 dapat disajikan seperti berikut:

Accuracy: 93%

Precision: 83.3%

Recall: 92.6%

F1-Score: 87.7%

3. Hasil dan Pembahasan

Halaman dashboard merupakan halaman pertama yang tampil setelah user melakukan login. Adapun implementasi dari halaman dashboard dapat dilihat pada Gambar 3.1 Halaman Dashboard.



Gambar 3.11 Halaman Dashboard

Halaman Prediksi merupakan halaman untuk memprediksi kelayakan. Adapun implementasi dari halaman prediksi dapat dilihat pada Gambar 3.2 Halaman Prediksi.

Gambar 3.2 Halaman Prediksi

Halaman hasil perbandingan Prediksi merupakan halaman untuk menampilkan hasil prediksi algoritma naïve bayes dan C4.5 dan disini dapat melihat score dari kedua algoritma. Adapun implementasi dari halaman hasil perbandingan prediksi dapat dilihat pada Gambar 3.3

Gambar 3.4 Halaman Hasil Perbandingan Prediksi

Halaman Sub Menu predictions merupakan halaman yang berisi data yang sudah di prediksi terlihat pada Gambar 3.4

Jumlah	Penghasilan	Prediksi	Status	Pekerjaan	Kondisi	Prediksi Naive Bayes	Prediksi C4.5	Score Naive Bayes	Score C4.5	Aksi
Keluarga Besar	Tinggi	SD	Bukit	Musikawan	Kurang	Tidak	Tidak	0	0	Detail
Keluarga Kecil	Kurang	SD	Kurang	petani	Kurang	Ya	Ya	0	0	Detail
Keluarga Besar	Kurang	SD	Kurang	petani	Kurang	Ya	Ya	0	0	Detail
Keluarga Kecil	Kurang	SD	Kurang	petani	Kurang	Ya	Ya	0	0	Detail
Keluarga Besar	Kurang	SD	Kurang	petani	Kurang	Ya	Ya	0	0	Detail
Keluarga Kecil	Kurang	SD	Kurang	petani	Kurang	Ya	Ya	0	0	Detail

Gambar 3.5 Halaman Sub Menu Data (Predictions).

Halaman Sub Menu Data Latih merupakan halaman yang berisi data latih dimana user dapat melakukan tambah, edit, hapus data dan melakukan import data berupa csv. Adapun implementasi dari halaman sub menu data latih dapat dilihat pada Gambar 3.5 Halaman Sub Menu Data (Data Latih).

NAMA	NIK	ALAMAT	JENJANG PENDIDIKAN	PENGINJILAN	STATUS	PEKERJAAN	KENDARAAN
ARI ANTONI	321024007070001	KPPAN BUKITINING RAS 004 RT 013	Keluarga Setengah	Kurang	TG	Sedang	Kurang
ARI ANTONI	321024008140004	KPPAN BUKITINING RAS 001 RT 003	Keluarga Setengah	Kurang	Tidak bekerja	Sedang	Kurang
ARI ANTONI	321024008100003	KPPAN BUKITINING RAS 002 RT 006	Keluarga Setengah	Kurang	SMP	Kurang	Kurang
ARI ANTONI	321024008000000	KPPAN BUKITINING RAS 001 RT 001	Keluarga Setengah	Tinggi	SMP	Baik	Baik
ARI ANTONI	321024008000000	KPPAN BUKITINING RAS 002 RT 004	Keluarga Setengah	Tinggi	DAS	Baik	Minicopa
ARI ANTONI	321024008100000	KPPAN BUKITINING RAS 004 RT 013	Keluarga Besar	Sedang	SMP	Sedang	Kurang

Gambar 3.6 Halaman Sub Menu Data (Latih).

Halaman Sub Menu Data Latih merupakan halaman yang berisi data latih dimana user dapat melakukan tambah, edit, hapus data dan melakukan import data berupa csv. Adapun implementasi dari halaman sub menu data latih dapat dilihat pada Gambar 3.6 Halaman Sub Menu Data (Data Latih).

Hasil pengujian algoritma naïve bayes

Naive Bayes Classification Results

Accuracy of Naive Bayes:98%

Confusion Matrix:

```
[[30  0]
 [ 1 29]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.97	1.00	0.98	30
1	1.00	0.97	0.98	30
accuracy			0.98	60
macro avg	0.98	0.98	0.98	60
weighted avg	0.98	0.98	0.98	60

Gambar 4. 2 Tabel Pengujian Sistem Berdasarkan Akurasi NB

Hasil pengujian algoritma C4.5

C4.5 (Decision Tree) Classification Results

Accuracy of C4.5: 0.9333333333333333

Confusion Matrix:

```
[[29 1]
 [ 3 27]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.91	0.97	0.94	30
1	0.96	0.90	0.93	30
accuracy			0.93	60
macro avg	0.94	0.93	0.93	60
weighted avg	0.94	0.93	0.93	60

Gambar 4. 3 Tabel Pengujian Sistem Berdasarkan Akurasi C4.5

4. Kesimpulan

Berdasarkan hasil penelitian dan laporan tugas akhir yang telah dilakukan, maka dapat ditarik kesimpulan bahwa Machine learning Penyortiran Buah Naga Berbasis Internet of Things Menggunakan Metode K-Means Pada Platform Blynk, maka penulis dapat menarik beberapa kesimpulan berdasarkan dari hasil penelitian sebagai berikut :

- Alat ini dapat membantu para petani buah dalam proses penyortiran.
- Sistem pendeteksi kualitas berat dan warna dapat berhasil terealisasi menggunakan sensor Load Cell untuk menghitung berat dan sensor TCS230 untuk menentukan warna. Dengan akurasi keberhasilan penimbangan sebesar 99% dan hanya memiliki error sebesar 1%. Serta sensor TCS3200 dapat mendeteksi warna dengan baik.
- Sistem penyortiran buah naga yang penulis kerjakan telah berfungsi dan berjalan baik sesuai dengan harapan.
- Data yang didapat oleh alat dapat dihitung menggunakan Algoritma K-means dengan benar.

Daftar Pustaka

- [1] R. W. Sinaga and R. Winanjaya, "Analisis Data Mining Menentukan Penerima Bantuan Langsung Tunai pada Desa Pamatang Purba dengan Algoritma C4.5," *Jurnal Teknologi dan Sistem Informasi*, vol. 3, no. 1, pp. 1–9, 2021.
- [2] G. Ramadhan et al., "Penerapan Data Mining Menggunakan Algoritma C4.5 Dalam Mengukur Tingkat Kepuasan Pasien BPJS," *Prosiding Seminar Nasional Riset dan Information Science (SENARIS)*, vol. 2, pp. 376–385, 2020.
- [3] L. Bachtiar and M. Mahradianur, "Analisis Data Mining Menggunakan Metode Algoritma C4.5 Menentukan Penerima Bantuan Langsung Tunai," *Jurnal Informatika*, vol. 10, no. 1, pp. 1–9, 2023.
- [4] Muhamad, M., Windarto, A. P., & Suhada, S. (2019). Penerapan Algoritma C4.5 Pada Klasifikasi Potensi Siswa Drop Out. KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer), 3(1), 1–8. <https://doi.org/10.30865/komik.v3i1.1688>
- [5] N. Rahmadani, R. Risnawati, and M. D. Sena, "Penerapan Algoritma Naïve Bayes dalam Penentuan Kelayakan Penerima Bantuan Program Keluarga Harapan," *Jurnal Teknisi*, vol. 3, no. 2, pp. 40–48, 2023.
- [6] R. S. Bhat and R. Bhat, "A Comparative Analysis of Machine Learning Algorithms for Classification Purpose," *Procedia Computer Science*, vol. 199, pp. 372–380, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050922021159>
- [7] S. Rasika, "Penerapan Algoritma C4.5 dalam Menentukan Faktor Kelayakan Penerima Bantuan Langsung Tunai (BLT-DD)," *Jurnal Ilmiah Tunas Bangsa*, vol. 1, no. 1, pp. 45–52, 2022.
- [8] J. R. Quinlan, "Improved Use of Continuous Attributes in C4.5," *Journal of Artificial Intelligence Research*, vol. 4, pp. 77–90, 1996. [Online]. Available: <https://jair.org/index.php/jair/article/view/10157>
- [9] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, 2009.
- [10] T. Wibowo, "Penerapan Data Mining Pemilihan Siswa Kelas Unggulan dengan Metode K-Means Clustering di SMP N 02 Tasikmadu," *Jurnal Ilmu Komputer*, vol. 5, no. 1, pp. 27–36, 2018.
- [11] A. Chandra, "Penerapan Data Mining Menggunakan Pohon Keputusan," *Jurnal Ilmu Komputer*, vol. 6, no. 3, pp. 1–6, 2017.