

# Penerapan Algoritma C4.5 Untuk Mengatur Persediaan Stok Barang Berbasis Website

Abastian Dwi Saputra<sup>1</sup>, Anita Qoiriah<sup>2</sup>

<sup>1,2</sup> Jurusan Teknik Informatika Fakultas Teknik Universitas Negeri Surabaya

<sup>1</sup>[abastian.18027@mhs.unesa.ac.id](mailto:abastian.18027@mhs.unesa.ac.id)

<sup>2</sup>[anitaqoiriah@unesa.ac.id](mailto:anitaqoiriah@unesa.ac.id)

**Abstrak**— Pandemi *Corona Virus Disease* (COVID-19) masih terjadi di Indonesia. Pencegahan COVID-19 yang dilakukan oleh pemerintah yaitu menerapkan peraturan Pembatasan Sosial Berskala Besar (PSBB) dan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM). Usaha Mikro, Kecil, dan Menengah (UMKM) adalah salah satu dampak dari diterapkan peraturan PSBB dan PPKM. Masalah yang terjadi pada UMKM yaitu mendapatkan masalah dalam menentukan stok persediaan barang, yang mana apabila barang mengalami kelebihan atau kekurangan akan menyebabkan kerugian. Dengan masalah yang terjadi pada UMKM, peneliti akan membantu UMKM dalam menentukan persediaan stok barang. Permasalahan tersebut dapat diatasi menggunakan *data mining* dengan klasifikasi algoritma C4.5. Algoritma C4.5 merupakan metode dengan mengklasifikasi data yang ada sehingga membentuk pohon keputusan. Pohon keputusan akan membantu menentukan persediaan stok barang yang perlu untuk ditambah atau tidak. Pada penelitian ini yaitu membantu UMKM penjual madu yang berjualan pada toko *online* untuk menentukan persediaan stok madu pada bulan berikutnya. Hasil dari penelitian dengan menggunakan algoritma C4.5 yaitu membentuk pohon keputusan yang dapat membantu UMKM dalam menentukan persediaan stok barang untuk satu bulan berikutnya. Algoritma C4.5 pada penelitian ini mempunyai *accuracy* sebesar 86,46%, *precision* sebesar 88,51%, *recall* sebesar 96,25%, dan F1 sebesar 92,22% .

**Kata Kunci**— UMKM, *Data Mining*, Algoritma C4.5, Persediaan Stok Barang, Prediksi Stok Barang.

## I. PENDAHULUAN

Pandemi *Corona Virus Disease* (COVID-19) di Indonesia masih terjadi di berbagai daerah hingga saat ini. Pencegahan COVID-19 telah dilakukan pemerintah dengan menerapkan peraturan yaitu Pembatasan Sosial Berskala Besar (PSBB) dan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM). Usaha Mikro, Kecil, dan Menengah (UMKM) adalah salah satu dampak dari PSBB dan PPKM.

Dampak pada UMKM akan memengaruhi penjualan karena UMKM kesulitan untuk menentukan persediaan stok barang. Permasalahan persediaan stok barang merupakan masalah dari dahulu hingga sekarang yang masih belum terselesaikan. Apabila UMKM mengalami kelebihan barang akan mengakibatkan banyak risiko seperti kemungkinan kerusakan barang, biaya perawatan barang, serta kerugian pada modal awal saat membeli barang. Selain itu apabila UMKM mengalami kekurangan barang maka akan memberikan dampak kehilangan keuntungan, kurang adanya kepercayaan yang diberikan oleh pelanggan sehingga pindah UMKM

lainnya, serta pelanggan akan kecewa terhadap UMKM kita [1].

Masalah dalam mengatur persediaan stok barang yang terjadi pada UMKM dapat diselesaikan menggunakan *data mining*. *Data mining* merupakan suatu teknik untuk mengumpulkan data hingga menemukan pola data dalam *dataset* yang berukuran besar, pola ini juga dapat memprediksi data baru [2]. *Data mining* yang cocok untuk diterapkan pada persediaan stok barang adalah klasifikasi menggunakan algoritma C4.5. Klasifikasi merupakan proses untuk membentuk suatu pola untuk memperkirakan data yang belum memiliki kelas dengan cara memisahkan kelas data satu dengan kelas data yang lainnya [3]. Algoritma C4.5 adalah metode yang digunakan untuk mengklasifikasikan data hingga membentuk suatu pohon keputusan (*Decision Tree*) [4]. *Decision tree* adalah suatu model yang dapat memprediksi suatu data sehingga membentuk pohon keputusan. *Decision tree* berfungsi sebagai alat pengambil keputusan untuk menemukan solusi dari masalah yang pernah terjadi sebelumnya [5].

Penelitian terkait dengan menerapkan metode algoritma C4.5 dalam kasus [6] memiliki latar belakang yaitu permasalahan pada stok barang apabila kekurangan akan menimbulkan kekecewaan pada pelanggan dan apabila stok barang kelebihan akan mengalami penumpukan stok barang. Pada kasus ini peneliti memberikan solusi dengan menggunakan metode algoritma C4.5 untuk memperkirakan waktu pembelian persediaan stok barang yang terjual dengan cara mengklasifikasi data barang mana yang akan distok ulang, sehingga ketersediaan stok barang tetap terjaga dan stabil.

Selain itu dalam kasus [7] memiliki latar belakang yaitu permasalahan pada pemasaran produk UMKM yang kurang laku sehingga mengelompokkan produk mana saja yang kurang laku saat dijual. Sehingga pada kasus ini peneliti menggunakan metode algoritma C4.5 untuk menentukan produk UMKM yang kurang laku dengan mengelompokkan indikator dalam membeli produk UMKM sehingga memberikan informasi strategi pemasaran yang akan dikeluarkan.

Selanjutnya dalam kasus [8] memiliki latar belakang yaitu permasalahan dalam menentukan penjualan barang. Sehingga pada kasus ini peneliti menggunakan algoritma C4.5 sebagai alat pendukung keputusan dengan cara membandingkan penjualan barang yang terjadi sebelumnya sehingga menemukan hasil data yang membentuk pohon keputusan.

Kemudian dalam kasus [9] memiliki latar belakang yaitu permasalahan pada perusahaan yang memerlukan cara untuk

memprediksi penjualan. Sehingga pada kasus ini peneliti menjelaskan bahwa penjualan memerlukan komputerisasi dengan didukung algoritma C4.5 sebagai alat untuk memprediksi penjualan sehingga dapat memudahkan melihat riwayat transaksi penjualan.

Terakhir dalam kasus [10] memiliki latar belakang yaitu permasalahan pada ketersediaan stok barang dengan memilih pemasok barang pada perusahaan. Sehingga pada kasus ini peneliti menggunakan algoritma C4.5 sebagai klasifikasi data yang berfungsi sebagai penentuan pemilihan pemasok barang dengan menggunakan salah satu atribut yaitu prediksi hasil penjualan produk pada periode berikutnya.

Menggunakan algoritma C4.5 dapat membantu dalam penelitian kali ini yaitu sebagai pendukung keputusan untuk menentukan stok barang. Selain itu, dapat juga sebagai alat prediksi penjualan yang akan datang sehingga dapat membantu menentukan jumlah stok barang pada bulan berikutnya [11]. Algoritma C4.5 juga memiliki akurasi yang tinggi dibanding algoritma lainnya yang telah diuji coba pada toko online yang sudah ada [12]. Algoritma C4.5 merupakan bagian dari *data mining* yang berfungsi untuk mengelompokkan data [13]. Data akan terus dikelompokkan hingga semua data mendapatkan kelompok [14]. Setelah data mendapatkan kelompok, maka data akan mendapatkan nilai. Kelompok dengan nilai tertinggi akan menjadi akar dari pohon keputusan [15].

UMKM membutuhkan sistem yang dapat mengakses informasi data produk dan mengatur persediaan stok barang di berbagai tempat, sistem tersebut adalah *website*. *Website* dapat memudahkan pemilik UMKM untuk mengakses informasi tentang produk dan dapat mengatur persediaan stok barang di berbagai tempat, sehingga pemilik UMKM tidak perlu menuju ke tempat UMKM hanya untuk mendapatkan informasi mengenai data produk dan untuk mengelola persediaan stok barang pada UMKM. Dengan adanya *website* pemilik UMKM dapat mengakses kapan pun dan di mana pun untuk mengatur persediaan stok barang.

Berdasarkan penjelasan di atas, maka penting bagi peneliti untuk membantu menyelesaikan permasalahan persediaan stok barang pada UMKM yang masih terjadi hingga sekarang. Untuk itu peneliti kali ini, menggunakan algoritma C4.5 yang berfungsi sebagai klasifikasi data dari sejumlah data yang sudah ada sebelumnya, sehingga dapat memperoleh suatu pola yang dapat membantu UMKM dalam menentukan persediaan stok barang yang perlu distok ulang. Selain itu UMKM juga perlu menggunakan *website* untuk mengatur dan mencatat persediaan stok barang. Dengan adanya penelitian ini UMKM mendapatkan kemudahan untuk mengetahui stok mana yang perlu distok ulang agar tidak salah dalam menentukan persediaan stok barang dan penelitian ini juga dapat membantu UMKM dalam mengambil keputusan untuk menentukan persediaan stok barang.

#### A. Algoritma C4.5

Algoritma C4.5 merupakan sebuah metode yang berfungsi untuk mengklasifikasikan data dari beberapa data

yang sudah ada sebelumnya. Setelah mendapatkan data, maka data tersebut akan dibagi sesuai dengan kriteria yang diinginkan. Data yang telah dibagi sesuai dengan kriterianya akan dikelompokkan.

Berikut cara untuk mengolah *data mining* dengan menggunakan algoritma C4.5:

1. Pilih *node*/akar  
Data yang ada akan disesuaikan dengan kelompok yang ada. Lalu menghitung nilai *Entropy* dan nilai *Gain* dari setiap kelompok. Nilai *Gain* yang tertinggi akan menjadi akar.
2. Tentukan kriteria untuk mengisi data.
3. Isi data sesuai dengan kriterianya.
4. Mengurangi perhitungan setiap kelompok sehingga semua data terisi.

Setelah semua data terisi pada kelompok masing-masing seperti langkah-langkah di atas, Selanjutnya yaitu menghitung nilai *Entropy* total dan menghitung nilai *Entropy* dari tiap kelompok yang ada. Kemudian menghitung nilai *Gain* dari tiap kelompok yang ada.

Rumus nilai *Entropy* dan nilai *Gain* dapat dilihat di bawah ini:

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

Dengan penjelasan  $\log_2$ :

$$\log_2(x) = \frac{\ln(x)}{\ln(2)}$$

Dengan:

S = Data contoh untuk training.

$p_i$  = proporsi yang didapat dari nilai *Gain*.

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

Dengan:

S = Data contoh untuk *training*.

A = Atribut/kelompok.

V = Nilai untuk atribut/kelompok.

Nilai(A) = Himpunan untuk atribut.

$|S_i|$  = Jumlah contoh data dari i.

$|S|$  = Jumlah seluruh contoh data.

$Entropy(S_i)$  = *Entropy* contoh data dari i.

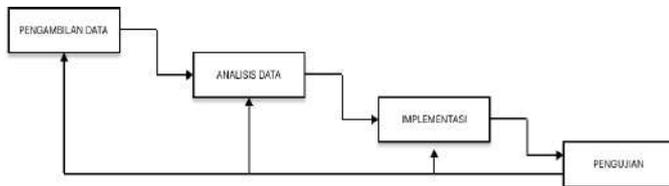
Secara umum algoritma C4.5 dapat menghasilkan pohon keputusan yang dapat dilihat sebagai berikut:

1. Menghitung *Entropy* total dari data yang ada, lalu dilanjutkan dengan *Entropy* masing-masing atribut.
2. Setelah memperoleh *Entropy* masing-masing atribut, lalu menghitung *Gain* masing-masing atribut.
3. Selanjutnya akar akan didapat dari kelompok dengan nilai *Gain* paling besar.

4. Kemudian untuk mendapatkan daun adalah dengan menghitung semua nilai Entropy dan nilai Gain dari tiap kelompok yang ada.

## II. METODOLOGI

Untuk kerangka kerja pada penelitian kali ini dengan melalui beberapa tahapan yang perlu dilakukan dari awal hingga akhir. Tahapan kerja tersebut dapat dilihat pada Gbr. 1.



Gbr. 1 Kerangka kerja penelitian.

### 1. Pengambilan Data

Pada tahap pengambilan data, peneliti melakukan wawancara kepada salah satu pemilik UMKM produk madu yang bernama *Sure Honey*. Hasil yang diperoleh dari wawancara adalah peneliti mendapatkan data nama produk yang dijual oleh pemilik UMKM pada toko *online*. Data tersebut dapat dilihat pada tabel I.

TABLE I  
DAFTAR NAMA PRODUK

Nama Produk
Botol Madu Nektar Bunga 200ml
Botol Madu Nektar Bunga 500ml
Sarang Lebah Madu Melifera 250g
Sarang Lebah Madu Melifera 500g
Jeriken Madu Nektar Bunga 500g
Jeriken Madu Nektar Bunga 1kg

Selain mendapatkan nama produk, peneliti juga mendapatkan data total stok dan data penjualan pada bulan sebelumnya. Total stok yang didapat adalah 120 dan total penjualan adalah 100.

*Dataset* yang akan diujikan dalam penelitian ini diperoleh dari hasil wawancara dengan salah satu UMKM produk madu bernama *Sure Honey*, hasil wawancara tersebut yaitu data total stok dan data penjualan pada satu bulan sebelumnya. Yang mana *dataset* tersebut akan memprediksi stok produk bulan berikutnya. *Dataset* tersebut dapat dilihat pada tabel II.

TABLE II  
DATA STOK DAN PENJUALAN

Nama Produk	Total Stok	Penjualan	Sisa Stok
Botol Madu Nektar Bunga 200ml	10	5	5
Botol Madu Nektar Bunga 500ml	10	1	9
Sarang Lebah Madu Melifera 250g	50	49	1
Sarang Lebah Madu Melifera 500g	30	25	5

Jeriken Madu Nektar Bunga 500g	10	10	0
Jeriken Madu Nektar Bunga 1kg	10	10	0

### 2. Analisa Data

Pada tahap analisa data, peneliti akan mengelompokkan produk. Untuk mengelompokkan produk, peneliti akan menggunakan *data mining*. Pada penelitian kali ini menggunakan data mining dengan klasifikasi algoritma. Algoritma C4.5 kali ini berfungsi untuk mengklasifikasi produk madu sesuai dengan kelompok yang ada.

Untuk mengelompokkan produk madu disini mempunyai 3 kelompok sebagai kriteria yang digunakan untuk memprediksi persediaan stok produk madu.

Pada penelitian ini, kelompok yang akan digunakan sebagai berikut:

- a. Nama Produk
- b. Bahan Produk
- c. Jenis Produk

Hasil akhir dengan menggunakan algoritma C4.5 yaitu membentuk pohon keputusan yang dapat membantu dalam memprediksi stok yang perlu ditambah atau tidak pada bulan berikutnya. Maka dari itu untuk mendapatkan hasil akhir tersebut, peneliti akan menggunakan *dataset* dari penjualan satu bulan sebelumnya yaitu terdapat total penjualan sebanyak 100. Detail penjualan satu bulan sebelumnya yang sudah dikelompokkan dapat dilihat pada tabel III.

TABLE III  
DETAIL PENJUALAN

Nama Produk	Bahan Produk	Jenis Produk
Botol Madu Nektar Bunga 200ml	Madu Nektar Bunga	Botol
Jeriken Madu Nektar Bunga 500g	Madu Nektar Bunga	Jeriken
Jeriken Madu Nektar Bunga 500g	Madu Nektar Bunga	Jeriken
Sarang Lebah Madu Melifera 500g	Madu Melifera	Sarang Lebah
Sarang Lebah Madu Melifera 500g	Madu Melifera	Sarang Lebah
Sarang Lebah Madu Melifera 500g	Madu Melifera	Sarang Lebah
Sarang Lebah Madu Melifera 500g	Madu Melifera	Sarang Lebah
Sarang Lebah Madu Melifera 500g	Madu Melifera	Sarang Lebah
Sarang Lebah Madu Melifera 250g	Madu Melifera	Sarang Lebah
Sarang Lebah Madu Melifera 250g	Madu Melifera	Sarang Lebah
Botol Madu Nektar Bunga 200ml	Madu Nektar Bunga	Botol
Jeriken Madu Nektar	Madu Nektar	Jeriken





Rancangan sistem pada Gbr. 2 akan dijelaskan sebagai berikut:

- a. Memasukkan data produk yang akan dikelompokkan berdasarkan data penjualan satu bulan sebelumnya.
  - b. Data produk yang telah dimasukkan akan dibagi berdasarkan kriteria yang telah ditentukan, kriteria tersebut yaitu nama produk, jenis produk, dan bahan produk.
  - c. Setelah menentukan kriteria yang diinginkan, selanjutnya memasukkan data berdasarkan kriteria yang telah ditentukan.
  - d. Setelah data mendapatkan kriteria masing-masing, selanjutnya data akan diolah sehingga menghasilkan nilai *Entropy*.
  - e. Mengecek kondisi apakah nilai *Entropy* sudah dihitung semua, jika belum maka sistem akan tetap menghitung perhitungan *Entropy* hingga selesai, jika sudah maka sistem akan menuju ke perhitungan *Gain*.
  - f. Menghitung nilai *Gain* berdasarkan nilai *Entropy* yang didapat. Nilai *Gain* terbesar akan menjadi akar dari kriteria.
  - g. Setelah nilai *Gain* didapat maka kriteria dengan nilai terbesar akan menjadi akar pohon keputusan. Kriteria tersebut akan menjadi penentu persediaan stok yang perlu distok ulang.
4. Implementasi
- Pada tahap implementasi, peneliti menggunakan algoritma C4.5 untuk melakukan perhitungan pada nilai *Entropy* dan nilai *Gain* dari data-data yang sudah dikelompokkan sehingga membentuk pola pohon keputusan.
5. Pengujian
- Pada tahap pengujian membandingkan data yang diperoleh sebelumnya dengan menggunakan algoritma C4.5 dengan hasil aplikasi berbasis *website*. Proses ini dapat membandingkan apakah ada kecocokan data yang dihasilkan antara menggunakan algoritma C4.5 dengan hasil yang didapat dengan menggunakan *website*.

### III. HASIL DAN PEMBAHASAN

#### A. Hasil

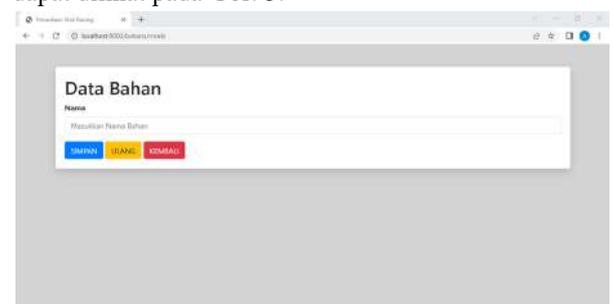
Pada tahap hasil, peneliti membuat *website* untuk memprediksi persediaan stok barang. Pada *website* yang akan dibuat, peneliti menggunakan algoritma C4.5 yang diterapkan untuk memprediksi persediaan stok barang pada bulan berikutnya.

1. Admin menuju halaman utama
- Pada tahap ini, admin akan melihat tampilan halaman utama dari *website* persediaan stok barang. Admin akan melihat berbagai menu yang dapat dipilih. Kali ini admin akan memilih menu data bahan. Tampilan halaman utama dapat dilihat pada Gbr. 3.

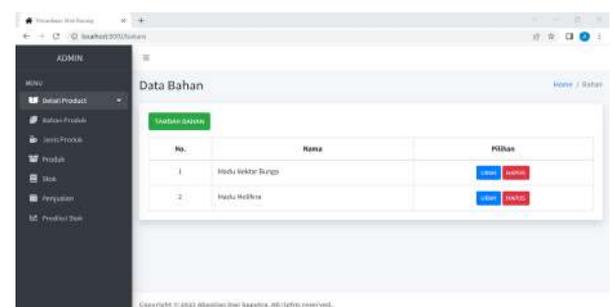


Gbr. 3 Admin menuju halaman utama.

2. Admin menambah data bahan
- Pada tahap ini, data bahan ini diperlukan untuk mengetahui produk madu ini terbuat dari bahan apa saja. Selanjutnya, admin dapat menekan tombol tambah bahan apabila ingin menambahkan data bahan. Berikutnya admin dapat mengisi kolom nama bahan dengan nama yang diinginkan. Untuk tampilan menambahkan data bahan dapat dilihat pada Gbr. 4. Setelah menambah data bahan, admin akan kembali menuju halaman data bahan untuk melihat semua daftar nama bahan yang telah ditambahkan. Berikutnya, admin dapat mengubah dan menghapus data bahan apabila diperlukan. Tampilan halaman data bahan dapat dilihat pada Gbr. 5.



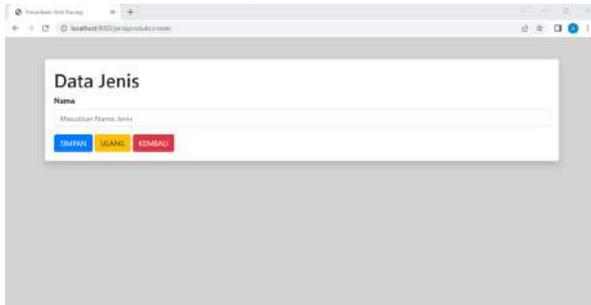
Gbr. 4 Admin menambah data bahan.



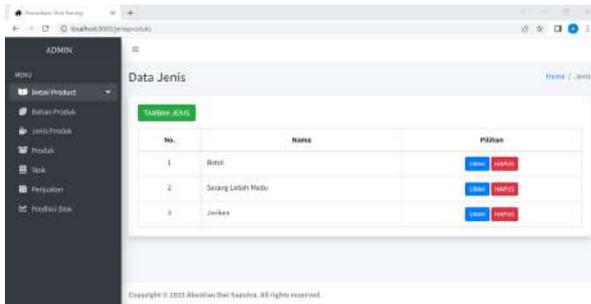
Gbr. 5 Admin menuju halaman data bahan.

3. Admin menambah data jenis
- Pada tahap ini sama seperti tahap data bahan, data jenis ini diperlukan untuk mengetahui produk madu ini dijual dalam bentuk jenis apa saja. Selanjutnya, admin dapat menekan tombol tambah jenis apabila ingin

menambahkan data jenis. Berikutnya admin dapat mengisi kolom nama jenis dengan nama yang diinginkan. Untuk tampilan menambahkan data jenis dapat dilihat pada Gbr. 6. Setelah menambahkan data jenis, admin akan kembali menuju halaman data jenis untuk melihat semua daftar nama jenis yang telah ditambahkan. Berikutnya, admin dapat mengubah dan menghapus data apabila diperlukan. Tampilan halaman data jenis dapat dilihat pada Gbr. 7.



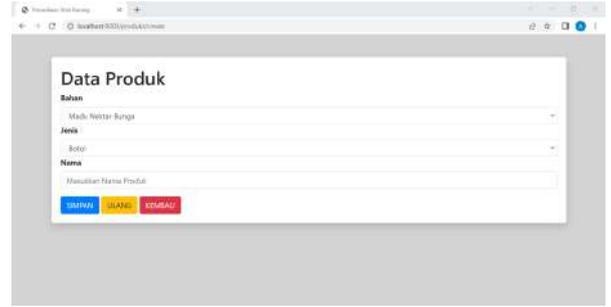
Gbr. 6 Admin menambah data jenis.



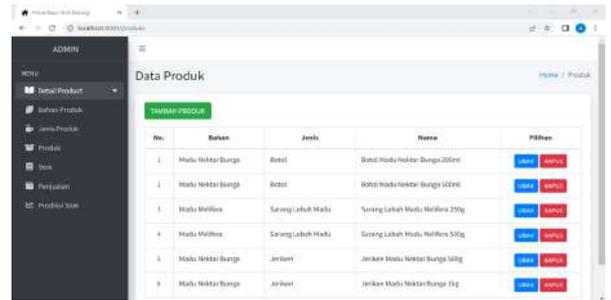
Gbr. 7 Admin menuju halaman data jenis.

#### 4. Admin menambah data produk

Pada tahap ini, data produk diperlukan untuk mengetahui produk madu apa saja yang ada berdasarkan bahan dan bentuk jenis yang telah dibuat pada tahap sebelumnya. Selanjutnya, admin dapat menekan tombol tambah produk apabila ingin menambahkan data produk. Berikutnya admin dapat memilih bahan yang diperlukan, jenis yang diperlukan, dan mengisi kolom nama produk dengan nama yang sesuai. Tampilan tambah produk dapat dilihat pada Gbr. 8. Setelah menambah data produk, admin akan menuju halaman produk untuk melihat semua daftar produk yang telah ditambahkan. Berikutnya, admin dapat mengubah dan menghapus data apabila diperlukan. Tampilan data produk dapat dilihat pada Gbr. 9.



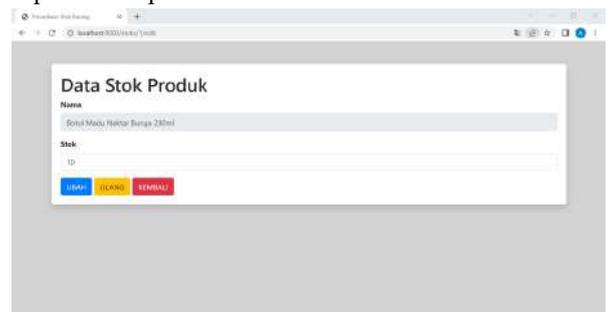
Gbr. 8 Admin menambah data produk.



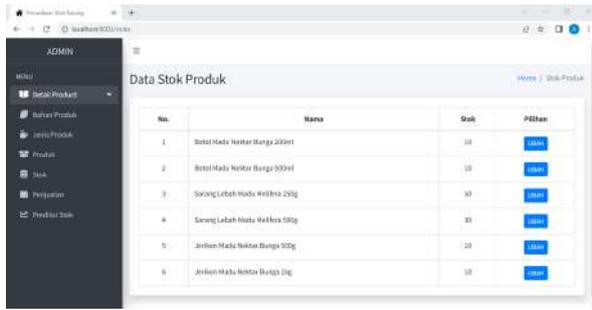
Gbr. 9 Admin menuju halaman data produk.

#### 5. Admin mengubah data stok produk

Pada tahap ini, admin dapat mengubah data stok produk berdasarkan stok produk madu satu bulan sebelumnya. Selanjutnya, admin dapat melihat daftar nama produk dan stok pada halaman data stok produk. Stok pada halaman ini didapat dari stok produk pada satu bulan sebelumnya. Berikutnya, admin dapat menekan tombol ubah untuk mengubah data stok produk. Admin dapat mengubah data stok berdasarkan nama produk sesuai dengan data stok satu bulan sebelumnya. Halaman ubah stok dapat dilihat pada Gbr. 10. Setelah mengubah data stok, admin akan menuju halaman data stok produk. Halaman data stok produk dapat dilihat pada Gbr. 11.



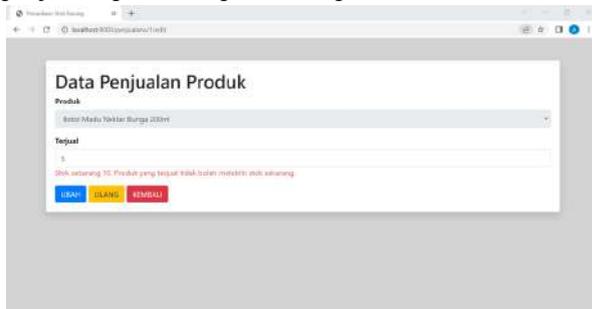
Gbr. 10 Admin mengubah data stok produk.



Gbr. 11 Admin menuju halaman data stok produk.

6. Admin mengubah data penjualan produk

Pada tahap ini, admin perlu mengubah data penjualan produk madu berdasarkan penjualan produk pada satu bulan sebelumnya. Selanjutnya, admin dapat melihat daftar nama produk dan penjualan pada halaman data penjualan produk. Penjualan pada halaman ini didapat dari hasil penjualan pada satu bulan sebelumnya. Berikutnya, admin dapat menekan tombol ubah untuk mengubah data penjualan produk. Admin dapat mengubah data penjualan berdasarkan nama produk sesuai dengan data penjualan satu bulan sebelumnya. Barang yang terjual tidak boleh lebih dari stok yang ada. Halaman penjualan dapat dilihat pada Gbr. 12. Setelah mengubah data penjualan, admin akan menuju halaman data penjualan produk. Halaman data penjualan produk dapat dilihat pada Gbr. 13.



Gbr. 12 Admin mengubah data penjualan produk.

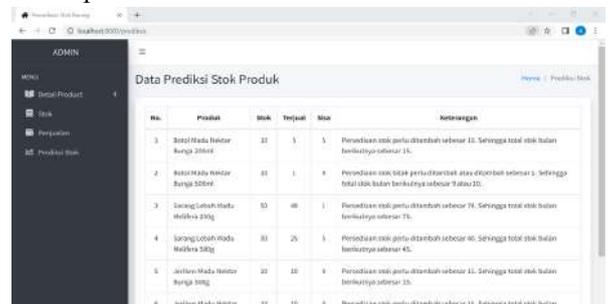


Gbr. 13 Admin menuju halaman data penjualan produk.

7. Admin melihat data prediksi stok produk

Pada tahap ini, admin dapat melihat data prediksi stok produk yang berisi nama produk, stok produk,

produk yang terjual, sisa produk, dan keterangan prediksi stok yang perlu atau tidak perlu untuk ditambah. Pada halaman ini, admin dapat melihat data prediksi stok yang perlu atau tidak perlu ditambah dengan melihat kolom keterangan. Prediksi ini menggunakan algoritma C4.5 yang didapat dengan menentukan stok pada bulan sebelumnya dan produk yang terjual pada bulan sebelumnya hingga menghasilkan sisa produk. Apabila sisa produk lebih dari atau sama dengan setengah total produk, maka akan menghasilkan keterangan persediaan stok perlu ditambah sebesar sisa stok ditambah dengan setengah total produk pada bulan sebelumnya. Sebaliknya apabila sisa produk kurang dari setengah total produk, maka akan menghasilkan keterangan persediaan stok tidak perlu ditambah atau ditambah sebesar produk yang terjual. Halaman data prediksi stok produk dapat dilihat pada Gbr. 14.



Gbr. 14 Admin menuju halaman data prediksi stok produk.

Dari hasil implementasi website yang sudah diujikan menghasilkan data prediksi stok produk dengan keterangan yang sesuai dengan yang diinginkan. Sehingga dapat membantu UMKM dalam menentukan persediaan stok barang untuk bulan berikutnya.

B. Pembahasan

Pada tahap pembahasan ini, peneliti akan menerapkan algoritma C4.5 pada data penjualan produk madu.

Data yang telah dikelompokkan berdasarkan kriteria yang telah ditentukan, selanjutnya diisi dengan nilai data yang berdasarkan dari total stok dan penjualan pada satu bulan sebelumnya. Data tersebut dapat dilihat pada tabel V.

TABLE V  
PENGELOMPOKAN DATA BESERTA NILAI

Node	Keterangan	Total Stok	Total Transaksi	
			Terjual	Sisa
1	Total produk	120	100	20
	Nama Produk	Botol Madu Nektar Bunga 200ml	5	5
		Botol Madu Nektar Bunga 500g	1	9

	Madu Nektar Bunga 500ml			
	Sarang Lebah Madu Melifera 250g	30	49	1
	Sarang Lebah Madu Melifera 500g	10	25	5
	Jeriken Madu Nektar Bunga 500g	10	10	0
	Jeriken Madu Nektar Bunga 1kg	10	10	0
Bahan Produk	Madu Nektar Bunga	40	26	14
	Madu Melifera	80	74	6
Jenis Produk	Botol	20	6	14
	Sarang Lebah Madu	80	74	6
	Jeriken	20	20	0

Setelah semua data yang ada telah dikelompokkan yang dapat dilihat pada tabel V, berikutnya adalah menghitung nilai *Entropy* total dan menghitung nilai *Entropy* dari tiap kelompok yang ada. Kemudian menghitung nilai *Gain* dari tiap atribut yang ada.

Berikut rumus untuk menghitung nilai *Entropy*:

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

Setelah mendapatkan rumus untuk menentukan nilai *Entropy* dan nilai *Gain*, Perhitungan nilai *Entropy* dan nilai *Gain* pada node/akar 1 adalah sebagai berikut::

a. Nilai *Entropy* Total

Pada perhitungan kali ini adalah menghitung nilai *Entropy* yang mewakili dari seluruh kelompok yang ada. Total stok pada penelitian ini adalah berjumlah 120 barang. Dengan rincian perhitungan 100 barang terjual dengan 20 sisa barang, dan jumlah stok barang semuanya adalah 120. Cara perhitungan sebagai berikut:

$$\begin{aligned} Entropy(total) &= \left(-\frac{100}{120} \times \log_2 \left(\frac{100}{120}\right)\right) + \left(-\frac{20}{120} \times \log_2 \left(\frac{20}{120}\right)\right) \\ &= 0,22 + 0,431 = 0,651 \end{aligned}$$

b. Nilai *Entropy* Nama Produk

Pada perhitungan *Entropy* kali ini terdapat 6 kategori yaitu Botol Madu Nektar Bunga 200ml, Botol Madu Nektar Bunga 500ml, Sarang Lebah Madu Melifera 250g, Sarang Lebah Madu Melifera 500g, Jeriken Madu Nektar Bunga 500g, dan Jeriken Madu Nektar Bunga 1kg. Perhitungan nilai *Entropy* sebagai berikut:

$$\begin{aligned} Entropy(Botol\ Madu\ Nektar\ Bunga\ 200ml) &= \left(-\frac{5}{10} \times \log_2 \left(\frac{5}{10}\right)\right) + \left(-\frac{5}{10} \times \log_2 \left(\frac{5}{10}\right)\right) \\ &= 0,5 + 0,5 = 1 \end{aligned}$$

$$\begin{aligned} Entropy(Botol\ Madu\ Nektar\ Bunga\ 500ml) &= \left(-\frac{1}{10} \times \log_2 \left(\frac{1}{10}\right)\right) + \left(-\frac{9}{10} \times \log_2 \left(\frac{9}{10}\right)\right) \\ &= 0,332 + 0,137 = 0,469 \end{aligned}$$

$$\begin{aligned} Entropy(Sarang\ Lebah\ Madu\ Melifera\ 250g) &= \left(-\frac{49}{50} \times \log_2 \left(\frac{49}{50}\right)\right) + \left(-\frac{1}{50} \times \log_2 \left(\frac{1}{50}\right)\right) \\ &= 0,028 + 0,113 = 0,141 \end{aligned}$$

$$\begin{aligned} Entropy(Sarang\ Lebah\ Madu\ Melifera\ 500g) &= \left(-\frac{25}{30} \times \log_2 \left(\frac{25}{30}\right)\right) + \left(-\frac{5}{30} \times \log_2 \left(\frac{5}{30}\right)\right) \\ &= 0,22 + 0,431 = 0,651 \end{aligned}$$

$$\begin{aligned} Entropy(Jeriken\ Madu\ Nektar\ Bunga\ 500g) &= \left(-\frac{10}{10} \times \log_2 \left(\frac{10}{10}\right)\right) + \left(-\frac{0}{10} \times \log_2 \left(\frac{0}{10}\right)\right) \\ &= 0 + 0 = 0 \end{aligned}$$

$$\begin{aligned} Entropy(Jeriken\ Madu\ Nektar\ Bunga\ 1kg) &= \left(-\frac{10}{10} \times \log_2 \left(\frac{10}{10}\right)\right) + \left(-\frac{0}{10} \times \log_2 \left(\frac{0}{10}\right)\right) \\ &= 0 + 0 = 0 \end{aligned}$$

c. Nilai *Entropy* Bahan Produk

Pada perhitungan *Entropy* kali ini terdapat 2 kategori yaitu Madu Nektar Bunga dan Madu Melifera. Perhitungan nilai *Entropy* sebagai berikut:

$$\begin{aligned} Entropy(Madu\ Nektar\ Bunga) &= \left(-\frac{26}{40} \times \log_2 \left(\frac{26}{40}\right)\right) + \left(-\frac{14}{40} \times \log_2 \left(\frac{14}{40}\right)\right) \\ &= 0,404 + 0,53 = 0,934 \end{aligned}$$

$$\begin{aligned} Entropy(Madu\ Melifera) &= \left(-\frac{74}{80} \times \log_2 \left(\frac{74}{80}\right)\right) + \left(-\frac{6}{80} \times \log_2 \left(\frac{6}{80}\right)\right) \\ &= 0,107 + 0,28 = 0,387 \end{aligned}$$

d. Nilai *Entropy* Jenis Produk

Pada perhitungan *Entropy* kali ini terdapat 3 kategori yaitu Botol, Sarang Lebah Madu, dan Jeriken. Perhitungan nilai *Entropy* sebagai berikut:

**Entropy(Botol)**

$$= \left(-\frac{6}{20} \times \log_2\left(\frac{6}{20}\right)\right) + \left(-\frac{14}{20} \times \log_2\left(\frac{14}{20}\right)\right)$$

$$= 0,521 + 0,36 = 0,881$$

**Entropy(Sarang Lebah Madu)**

$$= \left(-\frac{74}{80} \times \log_2\left(\frac{74}{80}\right)\right) + \left(-\frac{6}{80} \times \log_2\left(\frac{6}{80}\right)\right)$$

$$= 0,104 + 0,28 = 0,384$$

**Entropy(Jeriken)**

$$= \left(-\frac{20}{20} \times \log_2\left(\frac{20}{20}\right)\right) + \left(-\frac{0}{20} \times \log_2\left(\frac{0}{20}\right)\right)$$

$$= 0 + 0 = 0$$

e. Nilai *Gain* setiap atribut

Pada perhitungan *Gain* kali ini terdapat 3 kelompok yaitu Nama Produk, Bahan Produk, Jenis Produk. Perhitungan nilai *Gain* sebagai berikut:

**Gain(Nama Produk)**

$$= Entropy(s) - \sum_{i=1}^n \frac{|Nama\ Produk_i|}{Total} * Entropy(Nama\ Produk_i)$$

**Gain(Nama Produk)**

$$= 0,651 - \left(\left(\frac{10}{120}\right) \times 1 + \left(\frac{10}{120}\right) \times 0,469 + \left(\frac{50}{120}\right) \times 0,141 + \left(\frac{30}{120}\right) \times 0,651 + \left(\frac{10}{120}\right) \times 0\right)$$

$$= 0,651 - (0,08333 + 0,03908 + 0,05875 + 0,16275 + 0)$$

$$= 0,651 - 0,34391 = 0,30709$$

**Gain(Bahan Produk)**

$$= Entropy(s) - \sum_{i=1}^n \frac{|Bahan\ Produk_i|}{Total} * Entropy(Bahan\ Produk_i)$$

**Gain(Bahan Produk)**

$$= 0,651 - \left(\left(\frac{40}{120}\right) \times 0,934 + \left(\frac{80}{120}\right) \times 0,387\right)$$

$$= 0,651 - (0,31133 + 0,258)$$

$$= 0,651 - 0,56933 = 0,08167$$

**Gain(Jenis Produk)**

$$= Entropy(s) - \sum_{i=1}^n \frac{|Jenis\ Produk_i|}{Total} * Entropy(Jenis\ Produk_i)$$

**Gain(Jenis Produk)**

$$= 0,651 - \left(\left(\frac{20}{120}\right) \times 0,881 + \left(\frac{80}{120}\right) \times 0,384 + \left(\frac{20}{120}\right) \times 0\right)$$

$$= 0,651 - (0,14684 + 0,256 + 0)$$

$$= 0,651 - 0,40284 = 0,24816$$

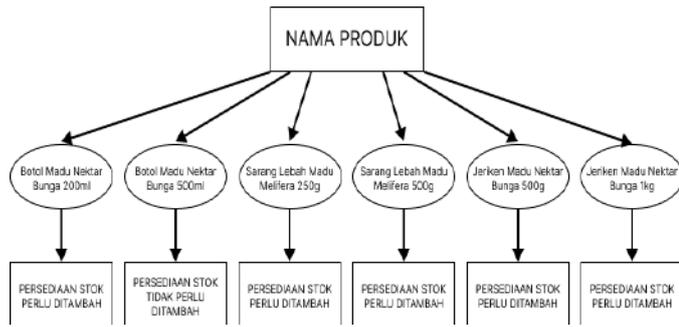
Hasil dari perhitungan nilai *Entropy* dan nilai *Gain* node 1 dapat dilihat pada tabel VI.

TABLE VI  
PERHITUNGAN *ENTROPY* DAN *GAIN*

Keterangan		Total Stok	Jumlah Penjualan		Entropy	Gain
			Terjual	Sisa		
Total Produk		120	100	20	0,651	
Nama Produk	Botol Madu Nektar Bunga 200ml	10	5	5	1	0,30709
	Botol Madu Nektar Bunga 500ml	10	1	9	0,469	
	Sarang Lebah Madu Melifera 250g	50	49	1	0,141	
	Sarang Lebah Madu Melifera 500g	30	25	5	0,651	
	Jeriken Madu Nektar Bunga 500g	10	10	0	0	
	Jeriken Madu Nektar Bunga 1kg	10	10	0	0	
Bahan Produk	Madu Nektar Bunga	40	26	14	0,934	0,08167
	Madu Melifera	80	74	6	0,387	
Jenis Produk	Botol	20	6	14	0,881	0,24816
	Sarang Lebah Madu	80	74	6	0,384	
	Jeriken	20	0	0	0	

Dari hasil perhitungan pada tabel VI, maka didapat nilai *Entropy* dan nilai *Gain*. Dapat diketahui bahwa nilai *Gain* tertinggi yaitu Nama Produk, maka dari itu Nama Produk akan menjadi akar dari pohon keputusan.

Berdasarkan perhitungan tersebut maka dapat didapat pohon keputusan. Maka didapat gambaran pohon keputusan seperti Gbr. 10.



Gbr. 15 Pohon keputusan hasil perhitungan algoritma C4.5.

Pada Gbr. 15 adalah hasil akhir dari proses *data mining* dengan klasifikasi menggunakan algoritma C4.5 yang dapat memprediksi stok produk dan juga dapat membantu dalam menentukan ketersediaan stok produk. Keterangan kriteria persediaan stok yang perlu ditambah adalah produk yang terjual lebih dari atau sama dengan setengah total stok produk. Persediaan stok yang perlu ditambah adalah hasil dari total produk yang terjual ditambah dengan setengah total stok produk pada bulan sebelumnya.

Hasil prediksi dari penelitian ini adalah sebagai berikut:

1. Jika Nama Produk Botol Madu Nektar Bunga 200ml persediaan stok perlu ditambah sebanyak 10.
2. Jika Nama Produk Botol Madu Nektar Bunga 500ml persediaan stok tidak perlu ditambah atau ditambah sebanyak 1.
3. Jika Nama Produk Sarang Lebah Madu Melifera 250g persediaan stok perlu ditambah sebanyak 74.
4. Jika Nama Produk Sarang Lebah Madu Melifera 500g persediaan stok perlu ditambah sebanyak 40.
5. Jika Nama Produk Jeriken Madu Nektar Bunga 500g persediaan stok perlu ditambah sebanyak 15.
6. Jika Nama Produk Jeriken Madu Nektar Bunga 1kg persediaan stok perlu ditambah sebanyak 15.

### C. Pengujian

Pada tahap ini, pengujian akan menghitung tingkat akurasi pada data tersebut menggunakan tabel *confusion matrix*. Dapat dilihat pada tabel VII.

TABLE VII  
CONFUSION MATRIX

		Nilai Sebenarnya	
		TRUE	FALSE
Nilai Prediksi	TRUE	TP (True Positive) Hasil yang benar	FP (False Negative) Hasil yang tidak diduga
	FALSE	FN (False Negative) Hasil yang hilang	TN (True Negative) Tidak ada hasil yang benar

- TP : Data positif terprediksi data positif
- FP : Data positif terprediksi data negatif
- FN : Data negatif terprediksi data positif
- TN : Data negatif terprediksi data negatif

Dari 120 data penjualan yang telah diuji sebelumnya (data *training*). Selanjutnya, akan dilakukan data *testing* dengan mengambil 80% atau sebanyak 96 data. Data *testing* merupakan data *training* yang terbaik digunakan untuk perbandingan. Data *testing* akan menggunakan pengukuran *confusion matrix*. Untuk pengukuran *confusion matrix* dapat dilihat pada tabel VIII.

TABLE VIII  
HASIL PENGUJIAN CONFUSION MATRIX

		Nilai Sebenarnya	
		TRUE	FALSE
Nilai Prediksi	TRUE	TP 77	FP 10
	FALSE	FN 3	TN 6

TP = Data testing yang didapatkan dari produk yang terjual dan terprediksi terjual.

FP = Data testing yang didapatkan dari produk terjual dan terprediksi tidak terjual.

FN = Data testing yang didapatkan dari produk yang tidak terjual dan terprediksi terjual.

TN = Data testing yang didapatkan dari produk yang tidak terjual dan terprediksi tidak terjual.

1. Hitung dengan menggunakan rumus *accuracy* sebagai berikut:

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

$$Accuracy = \frac{77 + 6}{77 + 6 + 10 + 3} \times 100\%$$

$$Accuracy = 86,46\%$$

2. Hitung dengan menggunakan rumus *precision* sebagai berikut:

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

$$Precision = \frac{77}{77 + 10} \times 100\%$$

$$Precision = 88,51\%$$

3. Hitung dengan menggunakan rumus *recall* sebagai berikut:

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

$$Recall = \frac{77}{77 + 3} \times 100\%$$

$$Recall = 96,25\%$$

4. Hitung dengan menggunakan rumus F1 sebagai berikut:

$$F1 = \frac{2 * P * R}{P + R} \times 100\%$$

$$F1 = \frac{2 * 0,8851 * 0,9625}{0,8851 + 0,9625} \times 100\%$$

$$F1 = 92,22\%$$

Dengan:

P = *Precision*

R = *Recall*

F1 = *F-measure*

Hasil yang diperoleh memperlihatkan *accuracy* sebesar 86,46%, *precision* sebesar 88,51%, *recall* sebesar 96,25%, dan F1 sebesar 92,22%.

#### IV. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang peneliti lakukan diperoleh kesimpulan sebagai berikut:

1. Algoritma C4.5 dapat diterapkan untuk mengelompokkan atribut. Atribut dikelompokkan berdasarkan produk yang ada.
2. Diperlukan *dataset* dari hasil Penjualan pada bulan sebelumnya sebagai acuan untuk memprediksi ketersediaan stok barang pada periode bulan berikutnya.
3. Dengan menggunakan rumus dari algoritma C4.5 yang dapat menghitung nilai *Entropy* dan nilai *Gain*, yang mana nilai *Gain* yang tertinggi dapat membantu menemukan akar dari pohon keputusan.
4. Pohon keputusan yang terbentuk dapat membantu sebagai alat prediksi stok yang perlu ditambah atau stok yang tidak perlu ditambah.
5. Hasil pengujian pada penelitian ini dengan membandingkan antara *website* dengan algoritma C4.5 memperoleh hasil *accuracy* sebesar 86,46%, *precision* sebesar 88,51%, *recall* sebesar 96,25%, dan F1 sebesar 92,22%.

#### V. SARAN

Untuk mengembangkan penelitian ini agar menjadi *website* yang lebih baik dan meningkatkan performa dari algoritma C4.5 sehingga ada beberapa saran dari peneliti yang bisa dikembangkan untuk penelitian berikutnya adalah sebagai berikut:

1. Penelitian ini dapat dibandingkan dengan metode *data mining* lainnya yang dapat membandingkan antara algoritma C4.5 dengan algoritma lainnya.
2. Perlu dilakukan penelitian lebih lanjut dengan kasus yang berbeda untuk mengetahui performa dari algoritma C4.5.
3. Website perlu dikembangkan seiring berkembangnya teknologi pada masa depan, karena website yang digunakan sangat sederhana dibandingkan dengan teknologi masa depan.

#### UCAPAN TERIMA KASIH

Ucapan syukur dan terima kasih peneliti sampaikan kepada:

1. Allah SWT atas segala rahmat dan hidayah, sehingga peneliti dapat menyelesaikan penelitian ini.
2. Bapak Adi Bambang Suhendro dan Ibu Rima Martin selaku kedua orang tua saya yang selalu mendoakan dan memberikan arahan serta motivasi untuk selalu memberikan yang terbaik hingga menyelesaikan penelitian.
3. Ibu Anita Qoiriah, S.Kom., M.Kom. selaku dosen pembimbing dengan sabar membimbing dan membantu memberikan saran dari awal hingga akhir penelitian.
4. Teman-teman yang selalu memberikan dukungan dan motivasi.
5. Kepada semua pihak yang telah membantu hingga selesainya penelitian ini.

#### REFERENSI

- [1] R. Rohim, I. Zufria, R. F. Nanuru, dan J. N. Anamofa, "C4.5 Classification Data Mining for Inventory Control," *International Journal of Engineering & Technology*, vol. 7, no. 2, pp. 68-72, 2018.
- [2] D. Nababan dan A. V. Tanlim, "Analisis Persediaan Stok Barang Menggunakan Algoritma C4.5 (Studi Kasus CV Harapan Jaya)," *Jurnal Ilmiah Informatika (JIF)*, vol. 7, no. 1, pp. 5-10, 2019.
- [3] A. K. Lalo, P. Batarius, dan Y. C. H. Siki, "Implementasi Algoritma C4.5 Untuk Klasifikasi Penjualan Barang di Swalayan Dutalia," *Jurnal Teknik Informatika Unika St. Thomas (JTIUST)*, vol. 6, no. 1, pp. 1-12, 2021.
- [4] Fadhila dan P. S. Hasugian, "Application of C4.5 Algorithm to Prediction Sales at PT Sumber Sayur Segar," *Journal of Intelligent Decision Support System (IDSS)*, vol. 5, no. 1, pp. 10-19, 2022.
- [5] A. H. Nasrullah, "Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris," *Jurnal Ilmiah Ilmu Komputer*, vol. 7, no. 2, pp. 45-51, 2021.
- [6] G. L. Pritalia, "Penerapan Algoritma C4.5 untuk Penentuan Ketersediaan Barang E-Commerce," *Indonesian Journal of Information Systems (IJIS)*, vol. 1, no. 1, pp. 47-56, 2018.
- [7] T. H. Zebua dan F. Riandari, "Implementation Of Data Mining With C4.5 Algorithm For Determining The Home Industry Product Marketing Strategy," *Journal of Intelligent Decision Support System (IDSS)*, vol. 4, no. 4, pp. 186-194, 2021.
- [8] Y. Rosela, "Implementasi Klasifikasi Decision Tree Menganalisa Status Penjualan Barang Menggunakan C4.5 (Studi Kasus: PT Matahari Department Store Medan Mall)," *Jurnal Pelita Informatika*, vol. 7, no. 3, pp. 404-411, 2019.
- [9] E. T. Putri, "Implementation of C4.5 Method and Artificial Neural Networks to Predict Sales," *Journal of Computer Science and Information Technology*, vol. 7, no. 4, pp. 110-115, 2021.

- [10] M. S. Abdulrahman dan J. A. Rasool, "Using Data Mining Algorithms to Predict Recommendations on Products," *International Journal of Advance Science and Technology*, vol. 29, no. 3, pp. 4370-4381, 2020.
- [11] A. Y. Ikhsanti, Y. Fauziah, dan R. I. Perwira, "Implementation of the C4.5 decision tree learning algorithm for sentiment analysis in e-commerce application reviews on google play store," *Computing and Information Processing Letters*, vol. 1, no. 1, pp. 25-30, 2021.
- [12] D. Alita, Setiawansyah, dan A. D. Putra, "C4.5 Algorithm for Motorcycle Sales Prediction On CV Mokus Rawajitu," *Journal of Sisfotek Global*, vol. 11, no. 2, pp. 127-134, 2021.
- [13] H. Lisnawati dan A. Sinaga, "Data Mining with Associated Methods to Predict Consumer Purchasing Patterns," *I. J. Modern Education and Computer Science*, vol. 5, pp. 16-28, 2020.
- [14] A. Massaro, A. Mustich, dan A. Galiano, "Decision Support System for Multistore Online Sales Based on Priority Rules and Data Mining," *Computer Science and Information Technology*, vol. 8, no. 1, pp. 1-12, 2020.
- [15] O. A. Alghanam, S. N. Al-Khatib, dan M. O. Hiari, "Data Mining Model for Predicting Customer Purchase Behavior in e-Commerce Context," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 13, no. 2, pp. 421-428, 2022.