

## PENERAPAN DATA MINING UNTUK MENGANALISA POLA PEMBELIAN SAYURAN HIDROPONIK MENGUNAKAN METODE ALGORITMA APRIORI

Sugeng Budi Rahardjo<sup>1,\*</sup>, Wiyanto<sup>2</sup>, Aprilia Sulistyohati<sup>3</sup>, Umilhuda<sup>4</sup>

<sup>1,2,4,\*</sup>Teknik Informatika, Universitas Pelita Bangsa

<sup>3</sup>Teknik Informatika, Universitas Indrapasta PGRI

<sup>1</sup>sugeng@pelitabangsa.ac.id, <sup>2</sup>wiyanto@pelitabangsa.ac.id, <sup>3</sup>aprilialia6891@gmail.com,

<sup>4</sup>umilhuda@gmail.com

### ABSTRAK

Seiring dengan maraknya usaha di bidang sayuran hidroponik, UD. Media Farms menyadari akan munculnya persaingan bisnis yang ketat, sehingga diperlukan strategi untuk menarik konsumen, kurangnya sumber informasi baru menjadi salah satu kendala kemajuan bisnis. Tujuan dari penelitian ini adalah untuk mengetahui informasi mengenai sayuran apa saja yang sering dibeli konsumen secara bersamaan dalam satu kali transaksi guna mengoptimalkan hasil panen sayuran hidroponik dan menentukan strategi bisnis penjualan sayuran hidroponik. Metode data mining yang digunakan dalam penelitian ini adalah metode association rule dengan algoritma apriori. Dari proses pengolahan data menggunakan algoritma apriori dan pengujian dengan alat rapid miner, dengan parameter support minimal 20% dan confidence minimal 60% maka aturan asosiasi yang terbentuk adalah 7 aturan. Dari 7 rule yang terbentuk, ditemukan 3 rule dengan nilai confidence tertinggi yaitu (selada → kangkung) dengan nilai support 50% dan confidence 79%, (selada, bayam merah → kangkung) dengan nilai support 25% dan kepercayaan diri 78%, (selada, pakcoy → kangkung) dengan nilai support 22% dan confidence 76%. Ketiga rule tersebut memiliki nilai lift ratio > 1 yang artinya rule tersebut valid dan dapat digunakan dengan baik.

**Kata kunci:** Penjualan sayuran hidroponik, *data mining*, *algoritma apriori*.

### ABSTRACT

Along with the rise of business in the field of hydroponic vegetables, UD. Media Farms is aware of the emergence of intense business competition, so a strategy is needed to attract consumers, the lack of new sources of information is one of the problems of business progress. The purpose of this study is to find out information about what vegetables are often purchased by consumers simultaneously in one transaction in order to optimize hydroponic vegetable yields and determine hydroponic vegetable sales business strategies. The data mining method used in this research is the association rule method with apriori algorithm. From the data processing process using the apriori algorithm and testing with rapid miner tools, with a minimum support parameter of 20% and a minimum confidence of 60%, the association rules formed are 7 rules. Of the 7 rules formed, 3 rules were found with the highest confidence values, namely (lettuce → kale) with a support value of 50% and confidence 79%, (lettuce, red spinach → kale) with a support value of 25% and confidence 78%, (lettuce, pakcoy → kale) with a support value of 22% and 76% confidence. The three rules have a lift ratio value > 1, which means that the rule is valid and can be used properly.

**Keywords:** Sales of hydroponic vegetables, *data mining*, *apriori algorithms*.

## PENDAHULUAN

Kebutuhan akan informasi yang akurat dan cepat sangat dibutuhkan dalam kehidupan sehari-hari seiring dengan perkembangan teknologi informasi. Namun, kebutuhan informasi yang meningkat untuk memperoleh pengetahuan baru terkadang tidak menyajikan informasi yang tepat, biasanya informasi tersebut masih perlu digali dari data dengan populasi yang cukup besar. Analisis data diperlukan untuk mengetahui potensi informasi yang ada. Hal ini mendorong munculnya bidang pengetahuan baru untuk memecahkan masalah penggalian informasi atau pola penting dari data dalam jumlah besar yang disebut data mining.

UD. Media Farms merupakan sebuah usaha yang bergerak dibidang pertanian (hortikultura) penanaman sayuran hidroponik, yang menyediakan kebutuhan sayuran hidroponik. Dengan kemajuan teknologi demi meminimalisir ketergantungan pada alam dan meningkatkan kualitas produk sayuran, UD. Media Farms telah memilih sistem hidroponik untuk membudidayakan sayuran yang sehat, segar, bersih dan berkualitas tinggi. UD. Media Farms memiliki lahan seluas 300 m<sup>2</sup> dengan total produksi 400-600 kg/Bulan, jenis komoditas yang ditanam yaitu, kangkung, pakcoy, selada, bayam merah, caisim, samhong dan selada keriting. Media Farms menjual hasil panennya ke beberapa restoran, supermarket dan masyarakat sekitar. Kapasitas produksi dan kualitas sayuran terus ditingkatkan untuk memenuhi permintaan pasar.

Seiring dengan maraknya bisnis di bidang sayuran hidroponik, UD. Media Farms menyadari akan timbulnya persaingan usaha yang ketat, sehingga diperlukan suatu strategi untuk menarik konsumen. Berbagai cara sudah ditempuh oleh pemilik usaha seperti memberikan promosi paket harga sayuran, menambah jenis sayuran yang ditanam dan memberikan diskon pada waktu tertentu. Namun, cara tersebut belum bisa menambah pendapatan secara optimal, bahkan menimbulkan terbuangnya sayuran yang kurang diminati konsumen. Hal ini disebabkan karena promosi yang dilakukan tidak memiliki acuan atau standar serta tidak berdasarkan data atau metode apapun sehingga terjadi ketidaksinkronan antara promosi dengan keinginan konsumen.

Terkait masalah tersebut diperlukan suatu cara untuk mengetahui pola hubungan jenis sayuran hidroponik yang sering dibeli secara bersamaan dalam satu transaksi yang dapat digunakan sebagai acuan atau standar dalam menentukan strategi bisnis. Dikarenakan terjadinya kegiatan penjualan setiap hari, data transaksi semakin lama semakin banyak. Namun pemilik usaha tidak mengetahui cara mengolah data untuk menemukan pola hubungan jenis sayuran yang sering dibeli secara bersamaan dalam satu transaksi. Selain itu, dalam mengolah data transaksi penjualan sayuran hidroponik masih manual, dimana data transaksi tersebut masih dicatat pada sebuah buku, sehingga jika data dibutuhkan akan memerlukan waktu yang lama untuk menemukannya. Dalam hal ini peran analisis dalam menganalisis data manual perlu dilengkapi dan diolah menjadi informasi yang berguna.

Maka pada penelitian ini penulis mencoba mengimplementasikan metode algoritma apriori dari data mining untuk menemukan informasi yang penting seperti pola pembelian sayuran oleh konsumen dengan keterkaitan item set yang terjual secara bersamaan.

## METODE PENELITIAN

### Data Mining

Data mining merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan menemukan informasi yang berguna dan pengetahuan yang terkait dari berbagai database besar. Data mining biasanya disebut sebagai knowledge discovery dalam database atau KDD adalah aktivitas yang menggunakan data historis untuk menemukan pola atau hubungan dalam sejumlah besar data. (Apridonal, Choiriah dan Akmal, 2019)

### Proses Data Mining

CRISP-DM (Cross Industry Standard Process for Data Mining) adalah standar yang dikembangkan pada tahun 1996 yang digunakan untuk proses analisis suatu industri sebagai strategi pemecahan masalah. CRISP-DM tidak menentukan standar atau karakteristik tertentu karena setiap data yang akan dianalisis akan diproses kembali pada tahap-tahap didalamnya. (Nofriansyah dan Nurcahuo, 2015),

Berikut tahap-tahap CRISP-DM

1. Pemahaman Bisnis (*Business Understanding*)  
Pada tahap ini meliputi pemahaman tujuan dan kebutuhan dari segi bisnis kemudian membuat strategi awal untuk mencapai tujuan.
2. Pemahaman Data (*Data Understanding*)  
Setelah menentukan tujuan bisnis, langkah selanjutnya adalah memahami data-data yang dibutuhkan dengan cara mengumpulkan data kemudian menentukan data apa saja yang akan digunakan untuk proses data mining.
3. Persiapan Data (*Data Preparation*)  
Dari data yang sudah dikumpulkan, lalu diolah sesuai dengan kebutuhan pada tahap pemodelan dengan cara memilih data, membersihkan data dan mengkonversi data.
4. Pemodelan (*Modeling*)  
Modeling adalah tahap menentukan algoritma dan tools yang akan dipakai untuk mendukung algoritma yang digunakan dan menentukan parameter yang disesuaikan untuk mendapat nilai optimal.
5. Evaluasi (*Evaluation*)  
Evaluasi merupakan tahap interpretasi hasil, pada tahap ini dilakukan evaluasi terhadap kualitas model apakah sudah memenuhi tujuan pada tahap pemahaman bisnis (*Business Understanding*).
6. Penyebaran (*Deployment*)  
Tahap *deployment* adalah tahap penyebaran hasil yang diperoleh dari proses data mining dengan membuat laporan agar dapat dimanfaatkan oleh pihak yang bersangkutan.

### **Association Rule**

Menurut Zhang yang dikutip oleh Edi Kurniawan, *association rule* adalah teknik untuk menemukan pola yang sering muncul didalam beberapa transaksi. Setiap transaksi terdiri dari beberapa item. Oleh karena itu, metode ini mendukung sistem rekomendasi dengan mencari pola setiap item dalam transaksi yang terjadi. (Kurniawan, 2019)

Metode dasar analisis asosiasi terbagi menjadi beberapa tahap, yaitu :

1. Pola Frekuensi Tinggi  
Pola frekuensi tinggi adalah tahap dimana pembentukan kombinasi antar item yang telah memenuhi syarat minimum *support* yang telah ditentukan sebelumnya. *Support* adalah nilai penunjang atau persentase kombinasi sebuah item dalam *database*, nilai *support* pada suatu item set dapat diperoleh dengan menggunakan sebuah rumus di bawah ini :

$$Support (A) = \frac{\sum \text{Transaksi yang mengandung } A}{\text{Total Transaksi}} \times 100$$

Rumus di atas berarti untuk menentukan nilai *support* pada satu item jumlah transaksi yang mengandung item A dibagi dengan jumlah transaksi yang ada pada database.

Sedangkan pada dua item set atau lebih diperoleh dengan rumus :

$$Support (A,B) = \frac{\sum \text{Transaksi yang mengandung } A \text{ dan } B}{\text{Total Transaksi}} \times 100\%$$

Rumus di atas digunakan untuk menentukan nilai *support* pada dua item set atau lebih, yaitu jumlah transaksi yang mengandung item A dan B dibagi dengan total transaksi yang terjadi pada data set.

2. Pembentukan Aturan Asosiasi  
Langkah selanjutnya setelah menentukan nilai *support* pada item set berfrekuensi tinggi lalu dibentuk aturan asosiasi (*confidence*) yang menyatakan kuatnya hubungan kombinasi item set pada transaksi. Untuk menentukan

aturan asosiasi yang terbentuk minimal item set harus memiliki dua kandidat A dan B. Untuk menentukan aturan  $A \rightarrow B$  digunakan rumus :

$$\text{Confidance } (A \rightarrow B) = \frac{\sum \text{Transaksi yang Mengandung A dan B}}{\sum \text{Transaksi yang Mengandung A}}$$

Untuk menentukan nilai *confidance* pada aturan asosiasi dilakukan dengan cara jumlah transaksi yang mengandung item A dan B dibagi dengan jumlah transaksi yang mengandung item A. 3. Join (Penggabungan) Join merupakan proses mengkombinasikan item dengan item yang lainnya sehingga tidak dapat terbentuk item lagi.

### 3. Prune (Pemangkasan)

Pemangkasan yaitu hasil dari kombinasi item yang tidak memenuhi syarat minimum support dipangkas untuk membuat kombinasi item berikutnya.

## Algoritma Apriory

Algoritma apriori adalah jenis aturan asosiasi pada data mining. Algoritma apriori digunakan agar komputer dapat mempelajari pola aturan asosiasi. Algoritma apriori digunakan untuk mencari kombinasi item set yang mempunyai suatu nilai keseringan tertentu sesuai kriteria yang ditentukan. [4]

Algoritma apriori merupakan algoritma paling terkenal untuk menemukan pola frekuensi tinggi. Pola frekuensi tinggi adalah pola-pola item di dalam suatu database atau support di atas ambang batas tertentu yang disebut dengan istilah minimum support yang memiliki frekuensi [5]

## Lift Ratio

*Lift ratio* adalah salah satu cara untuk menilai apakah aturan asosiasi itu kuat. *Lift ratio* biasanya digunakan sebagai penentu apakah aturan keakuratan suatu asosiasi. Nilai lift yang baik yaitu jika nilai lift  $> 1$  maka pada nilai ini proses transaksi dikatakan valid, karena keterkaitan item bergantung positif. Untuk menghitung lift ratio dapat menggunakan rumus berikut.[6]

$$\text{Lift Ratio} = \text{Benchmark Confidence } (A,B)$$

Untuk menghitung lift ratio sebelumnya harus mencari nilai benchmark confidence. Benchmark confidence merupakan jumlah perbandingan semua item yang menjadi consequent terhadap semua transaksi. Untuk mendapat nilai benchmark confidence dapat dihitung dengan rumus berikut.

Keterangan :

$N_c$  = Jumlah transaksi dengan item yang menjadi consequent

$N$  = Jumlah transaksi

## Prosedur Pengumpulan Data

Jenis data yang digunakan dalam penelitian ini adalah sebagai berikut.

### Data Primer

Data primer merupakan data yang dikumpulkan oleh perorangan atau suatu organisasi secara langsung dari objek yang diteliti. Pada penelitian ini penulis mendapatkan data primer dengan cara berikut :

#### 1. Observasi

Observasi dilakukan dengan cara mengamati secara langsung proses yang ada pada objek penelitian untuk mendapat data dan informasi yang real pada UD. Media Farms. Observasi dilakukan selama satu bulan yaitu sejak tanggal 08 Mei sampai dengan 08 Juni 2021.

#### 2. Wawancara

Penulis melakukan wawancara untuk mencari informasi mengenai penjualan sayuran hidroponik dengan mengajukan pertanyaan-pertanyaan yang mendukung permasalahan yang ada. Wawancara ini dilakukan dengan

tanya jawab secara langsung dengan pihak UD. Media Farms yaitu dengan bapak Dadang Media Laksana selaku pemilik UD. Media Farms, petani dan staff yang mengelola data transaksi.

### Data Sekunder

Data sekunder merupakan berbagai informasi yang telah ada sebelumnya yang digunakan untuk melengkapi kebutuhan data penelitian yang dengan sengaja dikumpulkan oleh peneliti.

Data sekunder yang digunakan dalam penelitian ini adalah sebagai berikut :

1. Data transaksi penjualan sayuran hidroponik dari bulan Juni 2020 – Mei 2021 pada UD. Media Farms yang berjumlah 1.283 data.
2. Buku dari Perpustakaan Digital *Library (iPusnas)* dan beberapa jurnal dari Google Scholar

## HASIL DAN PEMBAHASAN

### Perhitungan Algoritma Apriori

Dalam penelitian ini penulis menggunakan data transaksi penjualan sayuran hidroponik yang telah dilakukan proses pembersihan data, sehingga didapat 394 *data testing* yang akan digunakan untuk proses pemodelan.

### Analisa Pola Frekuensi Tinggi

Langkah awal dalam proses perhitungan algoritma apriori adalah menentukan batasan nilai *minimum support* yaitu sebesar 20%, kemudian dilakukan pencarian kombinasi semua jenis *item set*, kandidat yang tidak memenuhi syarat *minimum support* maka akan dipangkas.

Dari hasil seleksi kandidat 1 *item set* (C-1) digunakan untuk menentukan kandidat 2 *item set* (C-2) dan seterusnya. a.

### Pembentukan 1 Item Set

Proses pembentukan C1 atau disebut dengan 1 *item set* dengan jumlah *minimum support* = 20% dapat diselesaikan dengan rumus sebagai berikut :

$$Support (A) = \frac{\sum \text{Transaksi yang mengandung}}{\text{Total Transaksi}} \times 100\%$$

$$\begin{aligned} Support (\text{selada}) &= \frac{\sum \text{Transaksi yang mengandung selada}}{\text{Total Transaksi}} \times 100\% \\ &= \frac{248}{394} \times 100\% = 63\% \end{aligned}$$

$$\begin{aligned} Support (\text{kangkung}) &= \frac{\sum \text{Transaksi yang mengandung kangkung}}{\text{Total Transaksi}} \times 100\% \\ &= \frac{295}{394} \times 100\% = 75\% \end{aligned}$$

$$\begin{aligned} Support (\text{bayam merah}) &= \frac{\sum \text{Transaksi yang mengandung bayam merah}}{\text{Total Transaksi}} \times 100\% \\ &= \frac{227}{394} \times 100\% = 58\% \end{aligned}$$

$$\begin{aligned} \text{Support (pokcay)} &= \frac{\sum \text{Transaksi yang mengandung pokcay} \times 100\%}{\text{Total Transaksi}} \\ &= \frac{208 \times 100\%}{394} = 53\% \end{aligned}$$

$$\begin{aligned} \text{Support (caisim)} &= \frac{\sum \text{Transaksi yang mengandung caisim} \times 100\%}{\text{Total Transaksi}} \\ &= \frac{91 \times 100\%}{394} = 23\% \end{aligned}$$

$$\begin{aligned} \text{Support (samhong)} &= \frac{\sum \text{Transaksi yang mengandung samhong} \times 100\%}{\text{Total Transaksi}} \\ &= \frac{106 \times 100\%}{394} = 27\% \end{aligned}$$

$$\begin{aligned} \text{Support (selada keriting)} &= \frac{\sum \text{Transaksi yang mengandung selada keriting} \times 100\%}{\text{Total Transaksi}} \\ &= \frac{61 \times 100\%}{394} = 15\% \end{aligned}$$

Tabel 1. Kandidat 1 Item Set

<i>Item set</i>	<b>Jumlah</b>	<b>Support (%)</b>
Selada	248	63
Kangkung	295	75
Bayam merah	227	58
Pakcoy	208	54
Caisim	91	23
Samhong	106	27
Selada Keriting	61	15

Dari proses pembentukan 1 *item set* dengan *minimum support* 20%, maka *item set* yang memenuhi yaitu : selada, kangkung, bayam merah, pakcoy, caisim dan samhong. Maka untuk iterasi selanjutnya *item set* yang tidak memenuhi nilai *minimum support* akan dipangkas.

Tabel 2. Large 1 Item Set

<i>Item set</i>	<b>Jumlah</b>	<b>Support (%)</b>
Selada	248	63
Kangkung	295	75
Bayam merah	227	58
Pakcoy	208	53

Caisim	91	23
Samhong	106	27

### Kombinasi 2 Item Set

Setelah diketahui kandidat 1 *item set* yang memenuhi nilai *minimum support*. Selanjutnya pembedakan C2 atau disebut dengan 2 *item set* dengan jumlah *minimum support* = 20% dapat diselesaikan dengan rumus sebagai berikut:

$$Support(A,B) = \frac{\sum \text{Transaksi yang mengandung A dan B}}{\text{Total Transaksi}} * 100\%$$

Pada kombinasi 2 item set ini dapat dilihat pada table berikut :

Tabel 3. Kandidat 2 Item Set (C2-1)

Item set	Jumlah	Support (%)
Selada, kangkung	197	50
Selada, bayam merah	125	32
Selada, pakcoy	115	29
Selada, caisim	41	10
Selada, samhong	56	14
Kangkung, bayam merah	163	41
Kangkung, pakcoy	150	38

Tabel 4. Kandidat 2 Item Set (C2-2)

Item set	Jumlah	Support (%)
Kangkung, caisim	52	13
Kangkung, samhong	63	16
Bayam merah, pakcoy	80	20
Bayam merah, caisim	40	10
Bayam merah, samhong	47	12
Pakcoy, caisim	43	11
Pakcoy, samhong	46	12
Caisim, samhong	24	6

Berdasarkan hasil perhitungan nilai *support* pola kombinasi dua *item set* yang memenuhi *minimum support* 20% yaitu sebagai berikut :

Tabel 5. Large 2 Item Set (C2)

Item	Jumlah	Support (%)
Selada, kangkung	197	50
Selada, bayam merah	125	32
Selada, pakcoy	115	29
Kangkung, bayam merah	163	41
Kangkung, pakcoy	150	38
Bayam merah, Pakcoy	80	20

### Kombinasi 3 Item Set

Setelah diketahui kandidat 2 *item set* yang memenuhi nilai *minimum support*. Selanjutnya pembedakan C3 atau disebut dengan 3 *item set* dengan jumlah *minimum support* = 20% dapat diselesaikan dengan rumus sebagai berikut

$$Support(A,B,C) = \frac{\sum \text{Transaksi yang mengandung A, B dan C}}{\text{Total Transaksi}} * 100\%$$

Pada kombinasi 3 item set ini dapat dilihat pada table berikut :

Tabel 6. Kombinasi 3 Item Set (C3)

<i>Item set</i>	<b>Jumlah</b>	<b>Support (%)</b>
Selada, kangkung, bayam merah	98	25
Selada, kangkung, pakcoy	87	22
Selada, bayam merah, pakcoy	23	6
Kangkung, bayam merah, pakcoy	45	11

Berdasarkan hasil perhitungan nilai support pola kombinasi tiga item set, hanya ada satu kandidat yang memenuhi minimum support yang dapat dilihat pada tabel 7 di bawah ini.

Tabel 7. Large 3 Item Set

<i>Item set</i>	<b>Jumlah</b>	<b>Support (%)</b>
Selada, kangkung, bayam merah	98	25
Selada, kangkung, pakcoy	87	22

**Kombinasi 4 Item Set**

Pada iterasi pembentukan 3 item set masih didapatkan item set yang memenuhi nilai minimum support, maka dapat dilakukan iterasi pembentukan 4 item set.

Pembentukan C4 atau disebut dengan 4 item set dengan jumlah minimum support = 20% dapat diselesaikan dengan rumus sebagai berikut :

$$Support (A,B,C,D) = \frac{\sum Transaksi\ yang\ mengandung\ A,\ B,\ C\ dan\ D}{D} *$$

Pada kombinasi 4 item set ini dapat dilihat pada table berikut :

Tabel 8. Kombinasi 4 Item Set (C4)

<i>Item set</i>	<b>Jumlah</b>	<b>Support (%)</b>
Selada, kangkung, bayam merah, pakcoy	8	2

Berdasarkan iterasi pembentukan Gambar 2. Periksa Data4 item set, dapat dilihat bahwa hanya ada satu kandidat yang terbentuk dan kandidat tersebut tidak memenuhi nilai minimum support. Maka iterasi dihentikan hanya sampai pembentukan 3 item set saja. Hal ini merupakan ciri dari algoritma apriori

**Pembentukan Aturan Asosiasi**

Setelah pola frekuensi tinggi ditemukan, selanjutnya mencari aturan asosiasi yang memenuhi syarat minimum confidence dengan menghitung confidence aturan asosiatif A → B. Minimum confidence dalam penelitian ini ditentukan sebesar 60%. Mencari nilai confidence dapat menggunakan rumus berikut:

$$Confidence (A \rightarrow B) = \frac{\sum Transaksi\ yang\ mengandung\ A\ dan\ B}{\sum Transaksi} x$$

Pada pembentukan aturan asosiasi ini dapat dilihat pada table berikut :

Tabel 9. Hasil Perhitungan *Confidence*

<i>Rule</i>	<i>Confidence (%)</i>
Jika membeli selada maka akan membeli kangkung	79
Jika membeli kangkung maka akan membeli selada	67
Jika membeli selada maka akan membeli bayam merah	50
Jika membeli bayam merah maka akan membeli selada	55
Jika membeli selada maka akan membeli pakcoy	46
Jika membeli pakcoy maka akan membeli selada	55
Jika membeli kangkung maka akan membeli bayam merah	55
Jika membeli bayam merah maka akan membeli kangkung	72
Jika membeli kangkung maka akan membeli pakcoy	51
Jika membeli pakcoy maka akan membeli kangkung	72
Jika membeli bayam merah maka akan membeli pakcoy	35
Jika membeli pakcoy maka akan membeli bayam merah	38
Jika membeli selada dan kangkung maka akan membeli bayam merah	50
Jika membeli selada dan bayam merah maka akan membeli kangkung	78
Jika membeli kangkung dan bayam merah maka akan membeli selada	60
Jika membeli selada dan kangkung maka akan membeli pakcoy	44
Jika membeli selada dan pakcoy maka akan membeli kangkung	76
Jika membeli kangkung dan pakcoy maka akan membeli selada	58

Berdasarkan parameter yang telah ditentukan yaitu *minimum support* sebesar 20% dan *minimum confidence* sebesar 60% maka terbentuk empat aturan asosiasi yang dapat dilihat pada tabel berikut:

Tabel 10. Aturan Asosiasi Final

No	<i>Rule</i>	<i>Support (%)</i>	<i>Confidence (%)</i>
1	Jika membeli selada maka akanmembeli kangkung	50	79
2	Jika membeli kangkung maka akanmembeli selada	50	67
3	Jika membeli bayam merah maka akanmembeli kangkung	41	72
4	Jika membeli pakcoy maka akanmembeli kangkung	38	72
5	Jika membeli selada dan bayam merahmaka akan membeli kangkung	25	78
6	Jika membeli kangkung dan bayammerah maka akan membeli selada	25	60
7	Jika membeli selada dan pakcoy makaakan membeli kangkung	22	76

Setelah kombinasi *item set* sudah ditentukan, langkah terakhir adalah memvalidasi akurasi dari *rule* yang ditemukan yaitu dengan menghitung *lift ratio*. Perhitungan *lift ratio* dimulai dengan menghitung nilai *benchmark confidence*. *Benchmark confidence* adalah jumlah perbandingan semua *item* yang menjadi *consequent* terhadap semua transaksi.

Tabel 11. Perhitungan Benchmark Confidence

<i>Antecment</i>	<i>Consequent</i>	<i>Proses Benchmark Confidence</i>	<i>Benchmark Confidence (%)</i>
Selada	Kangkung	$(295/394)*100$	75
Kangkung	Selada	$(248/394)*100$	63
Bayam merah	Kangkung	$(295/394)*100$	75
Pakcoy	Kangkung	$(295/394)*100$	75
Selada, Bayam merah	Kangkung	$(295/394)*100$	75

Kangkung, BayamMerah	Selada	$(248/394)*100$	63
Selada, pakcoy	Kangkung	$(295/394)*100$	75

Tabel 12. Perhitungan *Lift Ratio*

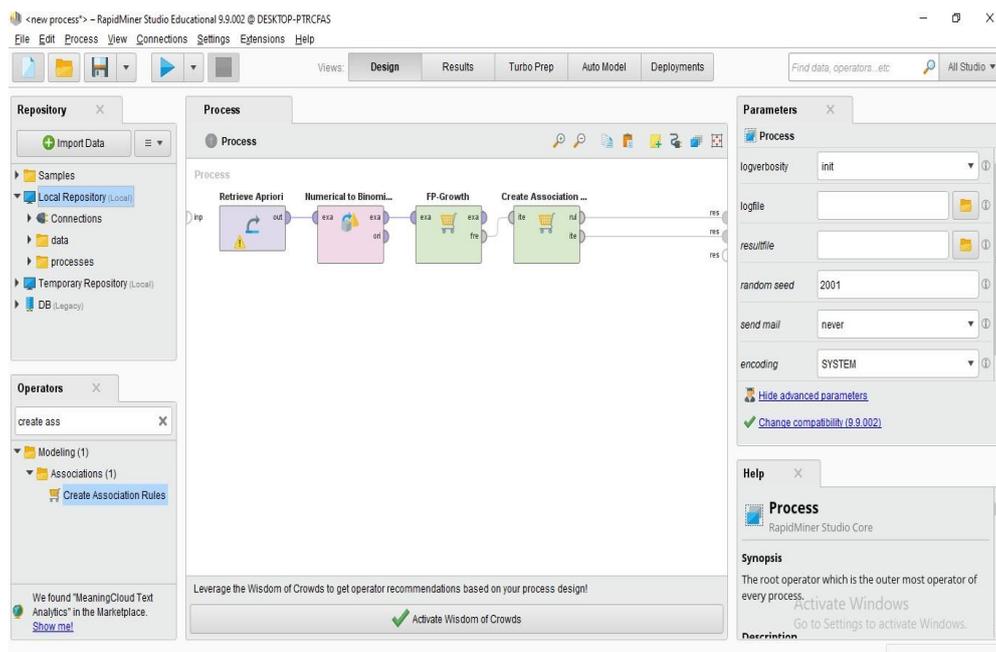
<i>Antecment</i>	<i>Consequent</i>	<i>Confidence (%)</i>	<i>Benchmark Confidence (%)</i>	<i>Lift Ratio</i>
Selada	Kangkung	79	75	1.060
Kangkung	Selada	67	63	1.060
Bayam merah	Kangkung	72	75	0.957
Pakcoy	Kangkung	72	75	0.961
Selada, Bayammerah	Kangkung	78	75	1.047
Kangkung, BayamMerah	Selada	60	63	0.955
Selada, Pakcoy	Kangkung	76	75	1.010

### Pengujian Dengan Tools *RapidMiner*

Setelah melakukan perhitungan manual algoritma apriori, selanjutnya dilakukan pengujian dengan tools RapidMiner. Berikut langkah-langkah pengujian dengan tools RapidMiner.

#### *Proses Algoritma Apriori*

Setelah data berhasil dimasukan, langkah selanjutnya yaitu tahapan proses algoritma apriori. Dengan mengklik menu *desain*, drag data ketampilanproses. Untuk mengolah data dengan metode algoritma apriori pada RapidMiner, memerlukan tiga operator. Operator inilah yang nantinya akan memproses data secara bertahap. Operator pertama yaitu *numerical to binominal* (untuk mengubah binary ke bilangan binominal) drag ke tampilanproses. Selanjutnya operator *FP-growth*, operator ini digunakan untuk menentukan *frequent item set* dengan mengisi parameter nilai *minimumsupport* sebesar 20% atau 0,2 drag ke tampilan proses. Kemudian operator *create association rule*, operator ini akan menghasilkan aturan asosiasi dari himpunan *frequent item set* yang diberikan, isi parameter nilai *minimum confidence* yang telah ditentukan. Pada pengujian kali ini *minimum confidence* yang ditentukan sebesar 60% atau 0,6 drag ketampilan proses. Hubungkan semua operator tersebut seperti gambar berikut.



Gambar 1. Tampilan Proses

### Hasil *Running* Algoritma Apriori

Untuk melihat hasil dari proses pengujian, klik *button play* untuk mengetahui hasil dari aturan asosiasi yang akan terbentuk. Maka tampilannya seperti berikut :

No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift	Conviction
1	Kangkung, Bayam Merah	Selada	0.249	0.601	0.883	-0.579	-0.012	0.955	0.929
2	Kangkung	Selada	0.497	0.664	0.856	-1	0.026	1.056	1.104
3	Bayam Merah	Kangkung	0.414	0.718	0.897	-0.739	-0.018	0.959	0.891
4	Pakcoy	Kangkung	0.381	0.721	0.904	-0.675	-0.015	0.963	0.901
5	Selada, Pakcoy	Kangkung	0.223	0.759	0.945	-0.365	0.003	1.013	1.041
6	Selada, Bayam Merah	Kangkung	0.249	0.772	0.944	-0.396	0.007	1.031	1.100
7	Selada	Kangkung	0.497	0.790	0.919	-0.761	0.026	1.056	1.198

Gambar 2. Tampilan Data Hasil Proses Pengujian

Setelah menjalankan proses untuk mencari *Association Rule* pada data set transaksi penjualan sayuran hidroponik. Algoritma *Apriori* dapat menghasilkan suatu aturan asosiasi. Dengan menentukan nilai minimum *support* sebesar 20%, Pada tahap pembentukan 1 item set terbentuk 6 item yang memenuhi syarat minimum *support*. keenam item tersebut digunakan untuk membentuk kombinasi 2 item set, kombinasi yang terbentuk sebanyak 6 kombinasi yang memenuhi syarat minimum *support*. Untuk pembentukan kombinasi 3 item set yang terbentuk ada 2 kombinasi item set yang memenuhi nilai minimum *support*.

Dari pembentukan kombinasi yang telah memenuhi nilai *minimum support*, selanjutnya dilakukan perhitungan nilai *confidence* dari masing-masing aturan, dengan menentukan *minimum confidence* sebesar 60%. Maka terbentuk 7 aturan asosiasi yang memenuhi nilai *minimum support* dan *minimum confidence*.

### KESIMPULAN

Berdasarkan hasil penelitian maka dapat diambil kesimpulan bahwa penerapan teknik data mining untuk menganalisa pola pembelian sayuran hidroponik dengan menggunakan metode algoritma apriori dapat diterapkan dengan baik dan mampu menghasilkan suatu aturan asosiasi dari data set transaksi penjualan yaitu berupa kombinasi jenis sayuran yang sering dibeli konsumen secara bersamaan dalam satu transaksi. Dari proses perhitungan algoritma apriori dengan parameter *minimum support* 20% dan *minimum confidence* 60%, aturan asosiasi yang terbentuk yaitu sebanyak 7 *rule*. Dari 7 *rule* yang terbentuk ditemukan 3 *rule* dengan nilai *confidence* tertinggi yaitu sebagai berikut:

1. Jika membeli selada maka akan membeli kangkung dengan nilai *support* 50% dan *confidence* 79%.
  2. Jika membeli selada dan bayam merah maka akan membeli kangkung dengan nilai *support* 25% dan *confidence* 78%.
  3. Jika membeli selada dan pakcoy maka akan membeli kangkung dengan nilai *support* 22% dan *confidence* 76%.
- Ketiga *rule* tersebut memiliki nilai *lift ratio* >1 yang berarti *rule* tersebut bernilai valid dan dapat dimanfaatkan dengan baik.

### SARAN

Untuk pengembangan penelitian selanjutnya, penulis memberikan beberapa saran sebagai berikut :

1. Penggunaan algoritma apriori sudah cukup baik, namun perlu dilakukan perbandingan dengan algoritma lain untuk menguji sejauh mana algoritma apriori dapat diandalkan dalam menemukan pola hubungan antar

*item.*

2. Dilakukan pengujian menggunakan berbagai *tools* untuk menganalisa hasil yang didapat agar mendapat hasil yang lebih valid

## DAFTAR PUSTAKA

- E. Kurniawan, (2019), "Implementasi Data Mining Dalam Analisa Pola Peminjaman Buku Di Perpustakaan Menggunakan Metode Association Rule," *Jurnal Teknologi dan Sistem Informasi*, vol. 5, no. 1, pp. 89–96.
- F. A. Sianturi, (2018), "Penerapan Algoritma Apriori Untuk Penentuan Tingkat Pesanan," *Mantik Penusa*, vol. 2, no. 1, pp. 50–57.
- M. Fauzy, K. R. Saleh W, and I. Asror, (2016), "Penerapan Metode Association Rule Menggunakan Algoritma Apriori pada Simulasi Prediksi Hujan Wilayah Kota Bandung," *J. Ilm. Teknol. Inf. Terap.*, vol. II, no. 2, pp. 221–227.
- Nofriansyah Dicky and G. W. Nurcahyo, (2019), *Algoritma Data Mining Dan Pengujian*. Yogyakarta: Deepublish.
- S. Nurajizah, (2019), "Analisa Transaksi Penjualan Obat menggunakan Algoritma Apriori," *INOVTEK Polbeng - Seri Inform.*, vol. 4, no. 1, p. 35.
- Y. Apridonan M, W. Choiriah, and A. Akmal, (2019), "Penerapan Data Mining Menggunakan Metode Association Rule Dengan Algoritma Apriori Untuk Analisa Pola Penjualan Barang," *Jurnal Teknologi dan Sistem Informasi*, vol. 5, no. 2, pp. 193–198.