

Memaksimalkan Potensi Penjualan Ikan Menggunakan Data Mining

Ilham Yori Pradana , Daninsyah Bagas Abiyansa , Harya Abimanyu Athalla A. Software Engineering

Dewi Rachmawati

Institut Teknologi Telkom Surabaya

Abstract—Indonesia adalah negara berkembang, yang memiliki wilayah kelautan yang luas dan juga negara kepulauan terbesar di dunia, Indonesia sendiri memiliki potensi sumber daya ikan yang sangat tinggi akan tetapi nelayan di Indonesia masih memiliki tingkat kemiskinan yang tinggi. Kehadiran program-program intervensi pembangunan, khususnya untuk masyarakat pesisir seperti program PEMP atau proyek Co-Fish, Departemen Kelautan dan Perikanan (DKP), merupakan wujud pengakuan adanya kemiskinan dikalangan masyarakat nelayan. Selain tingkat kemiskinan nelayan yang tinggi minat makan ikan masyarakat Indonesia masih sangatlah rendah dibandingkan dengan Singapura dan Malaysia. Oleh karena itu kementerian kelautan dan perikanan tengah menggalakan program GEMARIKAN, yang bertujuan untuk meningkatkan minat makan ikan masyarakat Indonesia dan juga untuk mengatasi Stunting pada anak. Untuk membantu program GEMARIKAN dan sekaligus untuk meningkatkan kesejahteraan nelayan, kami melakukan penelitian untuk memaksimalkan potensi penjualan ikan menggunakan *data mining*. *Data mining* merupakan proses Analisa data untuk menemukan suatu pola dari kumpulan data. Salah satu teknik *data mining* yang dapat digunakan adalah *association data mining* atau yang lebih dikenal *market basket analysis*. Tujuan dari metode ini untuk menemukan sejumlah *frequent itemset* dari data – data transaksi yang tersimpan dalam basis data. Dalam penelitian ini algoritma apriori digunakan untuk membantu menemukan sejumlah aturan asosiasi dari basis data transaksi penjualan ikan pada pasar ikan. Aturan tersebut dapat digunakan untuk melihat jenis ikan apa saja yang memiliki keterikatan tinggi saat terjual. Sehingga dapat meningkatkan kemungkinan pembeli akan membeli ikan-ikan tersebut bersamaan. Di harapkan dengan cara tersebut penjualan ikan dapat meningkat.

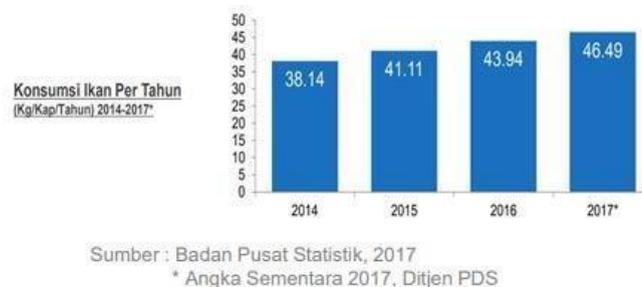
Keywords :*Data mining, itemset, market basket analysis, penjualan ikan, transaksi ikan.*

A. Pendahuluan

Indonesia adalah negara berkembang, yang memiliki wilayah kelautan yang luas dan juga negara kepulauan terbesar di dunia, Indonesia sendiri memiliki potensi sumber daya ikan yang sangat tinggi akan tetapi nelayan di Indonesia masih memiliki tingkat kemiskinan yang tinggi. Kehadiran program-program intervensi pembangunan, khususnya untuk masyarakat pesisir seperti program PEMP atau proyek Co-Fish, Departemen Kelautan dan Perikanan (DKP), merupakan wujud pengakuan adanya kemiskinan di kalangan masyarakat nelayan.

Selain tingkat kemiskinan nelayan yang tinggi minat makan ikan masyarakat Indonesia masih sangatlah rendah dibandingkan dengan Singapura dan Malaysia. Oleh karena itu kementerian kelautan dan perikanan tengah menggalakan program GEMARIKAN, yang bertujuan untuk meningkatkan minat makan ikan masyarakat Indonesia dan juga untuk mengatasi Stunting pada anak.

Program GEMARIKAN sendiri sudah diluncurkan sejak tahun 2004 oleh presiden RI, dan terus berkembang hingga saat ini. Dampak dari program tersebut terhadap masyarakat Indonesia dapat dilihat pada gambar dibawah pada tahun 2014 hingga 2017 konsumsi ikan masyarakat Indonesia mulai meningkat dari tahun ke tahun.



Gambar I-1. Konsumsi Makanan Jenis Ikan Per Tahun

Meningkatnya minat konsumsi makan ikan masyarakat Indonesia diharapkan juga dapat berdampak pada peningkatan penjualan ikan para nelayan. Jika potensi peningkatan penjualan ikan dimaksimalkan hal ini dapat membantu pemerintah dalam mengatasi rendahnya tingkat kesejahteraan nelayan. Untuk memaksimalkan potensi transaksi banyak cara yang dapat digunakan, salah satunya adalah *data mining*. Metode ini sering digunakan untuk mencari suatu pola yang nantinya akan digunakan untuk mengambil keputusan.

Diharapkan hasil dari penerapan *data mining* ini, dapat meningkatkan penghasilan dari nelayan dan membantu pemerintah untuk mengatasi kemiskinan pada nelayan.

A. Rumusan Masalah

Bagaimana cara meningkatkan potensi penjualan ikan menggunakan *data mining*? Apakah penerapan *data mining* kedalam data transaksi ikan dapat meningkatkan potensi penjualan?

B. Tujuan & Manfaat

Tujuan dari penelitian adalah :

- Melihat dampak dari penerapan metode *data mining* pada transaksi penjualan ikan.
- Menganalisis kinerja dari metode *Market Basket Analysis* dalam transaksi penjualan ikan.

Manfaat dari penelitian adalah :

- Membantu para penjual ikan, agar dapat memaksimalkan hasil penjualan yang mereka dapatkan.

C. Hipotesis

Berdasarkan analisis yang akan diteliti pola yang dihasilkan dari metode *Market Basket Analysis* akan dipengaruhi dari kebiasaan konsumen dalam melakukan pembelian, dari pola yang dihasilkan tersebut para penjual ikan dapat melihat item mana sajakah yang memiliki keterikatan tinggi, item-item yang memiliki keterikatan tinggi ini nantinya akan digunakan untuk mempengaruhi konsumen untuk membeli dua / lebih item secara bersamaan. Setelah pola ditemukan nantinya metode *Hotelling* akan diterapkan kedalam pola tersebut.

D. Rancangan Penelitian

Dalam penelitian yang menggunakan *Market Basket Analysis* & *Hotelling*, kami akan melakukan pengolahan data transaksi ikan terlebih dahulu yang kemudian dilanjutkan dengan penerapan metode *Hotelling*. Berikut ini rancangan penelitian dari kelompok kami :

a. Pengumpulan Data

Pada tahap ini kita menentukan variable apa saja yang akan digunakan kedalam penelitian ini. Setelah variable ditentukan kita mencari data yang berkaitan dengan variable tersebut. Seperti data transaksi ikan berdasarkan jenis ikan.

b. Analisis Data

Pada tahap ini, data yang kita dapat nantinya akan dilakukan *data cleaning* sehingga data yang tidak

konsisten tidak mempengaruhi hasil, setelah itu dilanjutkan dengan penerapan rule untuk mencari pola dari data yang ada.

c. *Menganalisa Hasil*

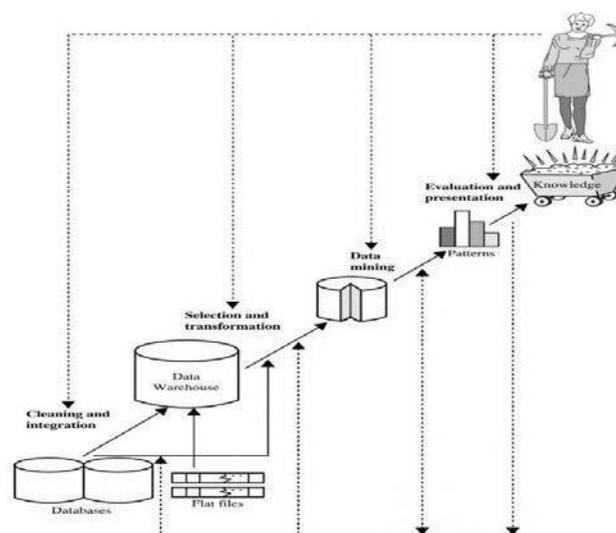
Pada tahap ini, pola yang tersedia akan di filter untuk menemukan pola yang terbaik dan akan dilanjutkan dengan penerapan metode hotelling kedalam pola tersebut. Hasil dari penerapan metode hotelling akan dianalisa dan dilanjutkan dengan penarikan kesimpulan apakah penerapan metode hotelling dan market basket analysis dapat meningkatkan penjualan ikan.

E. *Tinjauan Pustaka*

1) *Data mining*

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam basis data. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar [1]. Pakar lain berpendapat bahwa, “*Data mining* adalah proses menemukan suatu pola yang menarik dan pengetahuan dari sejumlah data.” [2]. *Data mining* sendiri memiliki beberapa jenis. Mereka dibagi menjadi beberapa kelompok berdasarkan tugas/pekerjaan yang dapat dilakukan, yaitu deskripsi, estimasi, prediksi, klasifikasi, pengklusteran, dan aturan asosiasi [3].

Proses yang harus dilalui untuk mendapatkan suatu informasi dapat digambarkan sebagai proses dibawah :



Gambar I-1. Gambar Proses Informasi

1. Data Cleaning. Didalam proses ini data akan dipilah yang tujuannya agar data yang tidak konsisten tidak masuk kedalam proses.
2. Data Integration. Proses penggabungan data yang telah diseleksi yang nantinya akan disimpan kedalam database.
3. Data Selection. Penyeleksian data yang relevan dengan penelitian yang diambil dari database.
4. Data Transformation. Proses transformasi dan konsolidasi data ke dalam berbagai bentuk yang sesuai untuk *data mining*.
5. Data Mining. Proses mengekstraksi data untuk menemukan sebuah pola, menggunakan metode-metode tertentu.
6. Pattern Evaluation. Mengidentifikasi pola yang menarik dan yang mewakili informasi dari data.
7. Knowledge. Informasi yang telah jadi yang nantinya akan ditampilkan kepada pengguna dengan cara yang menarik [2].

Dalam penelitian ini, kita menggunakan aturan asosiasi. Aturan asosiasi atau yang lebih dikenal sebagai *market basket analysis* adalah teknik *data mining* untuk menemukan suatu kombinasi antar item yang memiliki keterikatan tinggi, yang nantinya akan menghasilkan suatu itemset. Itemset ini digunakan bertujuan untuk meningkatkan kemungkinan pelanggan dapat membeli item yang memiliki keterikatan tinggi secara bersamaan [3]. *Market basket analysis* telah diterapkan pada berbagai bidang, seperti marketing, perbankan, data medis, data kriminalitas, dan data satelit [4].

“Market basket analysis memiliki suatu konsekuensi yang perlu diperhatikan dalam penggunaannya, yaitu semakin banyak produk barang yang terdapat di database akan menyebabkan rules yang dihasilkan oleh market basket analysis semakin meningkat. Dari sekian banyak rules tersebut, hanya beberapa rules saja yang bersifat informatif dan mampu mengungkap pola transaksi yang sebelumnya tidak ada. Berdasarkan hal ini, maka diperlukan suatu proses pemilihan rules yang didasarkan pada suatu ukuran kebaikan. Ukuran kebaikan yang umumnya digunakan untuk pemilihan rules adalah support, confidence, dan lift. Semakin besar ketiga nilai tersebut, maka rules tersebut semakin baik dalam memprediksi pola pembelian konsumen.” [5].

“Suatu rule dikatakan menarik dan memiliki kemampuan prediksi yang baik apabila rule ini memiliki nilai ukuran kebaikan yang tinggi dalam database. Tiga ukuran kebaikan yang pada umumnya digunakan

adalah support, confidence, dan lift. Ukuran kebaikan support mengukur proporsi kemunculan item dalam transaksi. Secara matematis, support dituliskan sebagai :

$$\text{Support}(A \rightarrow B) = \frac{\text{Count}(A \cap B)}{N}$$

Dengan N merupakan total banyaknya transaksi di dalam database. Ukuran kebaikan confidence mengukur kemampuan prediksi atau tingkat kepercayaan dari suatu rule. Secara matematis confidence dituliskan sebagai :

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Support}(A \rightarrow B)}{\text{Support}(A)}$$

Ukuran kebaikan lift mengukur besarnya peluang item B muncul apabila diberikan item A. Secara matematis nilai lift dihitung dengan menggunakan rumus :

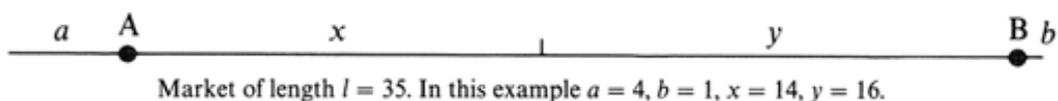
$$\text{Lift}(A \rightarrow B) = \frac{\text{Support}(A \rightarrow B)}{\text{Support}(A) \text{Support}(B)}$$

Suatu rule dikatakan baik apabila memiliki nilai support dan confidence yang mendekati 1 (atau 100%) dan lift di atas nilai 1." [5].

Metode Market Basket Analysis ini sering digunakan pada penelitian-penelitian terdahulu untuk melakukan pengambilan keputusan dalam bidang pemasaran. Salah satunya penelitian yang dilakukan oleh M. Hemalatha pada tahun 2012, penelitian tersebut menggunakan Market Basket Analysis untuk pasar ritel yang berada di India. Pada penelitian tersebut Market Basket Analysis digabungkan dengan menggabungkan K-Means dan predictive profiling yang digunakan untuk memaksimalkan potensi informasi yang didapatkan dan juga memprediksi sifat pelanggan di masa depan.

Hotelling Law

Hotelling Law adalah teori yang digunakan untuk suatu pedagang atau perusahaan untuk memilih tempat mereka berjualan, dalam teori ini ada dua kompetitor (duopoly), mereka menjual produk yang sama dan jarak antara dua kompetitor itu digambarkan seperti gambar berikut.



A dan B adalah kompetitor dan pembeli akan disebar pada garis l diatas. Harga antara dua kompetitor dinyatakan dalam p_1 dan p_2 , dan untuk jumlah yang terjual akan

dinyatakan sebagai qq_1 dan qq_2 . Kedua kompetitor itu akan bersaing untuk menarik pelanggan akan tetapi agar kompetitor B dapat menarik pelanggan pada bagian b dan y , maka harga / PP_2 dari B tidak boleh melebihi A / PP_1 dan

biaya transportasi pelanggan menuju ke A dari B sehingga dapat dinyatakan $P_2 = P_1 - c$. Maka keseimbangan harga akan terjadi antar kompetitor. Sehingga dapat dinyatakan sebagai berikut $PP_2 + cccc = PP_1 + cccc$ [6].

Metode ini sering sekali digunakan untuk memaksimalkan potensi penjualan antar kompetitor. Salah satu penelitian terdahulu yang pernah dilakukan oleh Alberto A. Pinto & Telmo Parreira pada tahun 2014, penelitian tersebut menggunakan konsep metode hotelling tetapi harga antar kompetitor tidak pasti, dalam penelitian tersebut membuktikan bahwa formula yang mereka terapkan tidak tergantung pada distribusi biaya produksi perusahaan.

I. METODE

Penelitian ini menggunakan dataset online retail yang berada di internet. Kemudian, data tersebut dianalisa dan dihitung menggunakan metode *market basket analysis* untuk menemukan pola dari data yang ada, program yang digunakan adalah MLxtend. Setelah pola ditemukan akan dilanjutkan penerapan metode hotelling.

A. Pengumpulan Data

Penelitian yang akan dilakukan menggunakan data dummy yang berasal dari *online-retail*, data tersebut berjumlah 541.910. Diharapkan untuk penelitian selanjutnya data yang digunakan adalah data transaksi asli dan untuk variabelnya menggunakan jenis ikan. Pengambilan data tersebut dapat menerapkan beberapa teknik antara lain :

- 1) *Studi Literatur* : Dalam proses ini peneliti diharuskan untuk mempelajari dan memahami tentang metode market basket analysis dan hotelling, sehingga peneliti mengetahui metode yang tepat untuk diterapkan dengan jumlah data yang tersedia.
- 2) *Wawancara* : Pada proses ini peneliti akan melakukan wawancara terhadap penjual untuk mengetahui berapa banyak ikan yang terjual, dan pembeli untuk mengetahui ikan apa yang dibeli.
- 3) *Observasi* : Proses ini dilakukan untuk memastikan apakah data yang didapat adalah data yang akurat & terpercaya. Selain itu untuk mengetahui apakah ada variable lain yang dapat mempengaruhi data.

B. Analisis Data

Setelah data didapatkan, proses penghitungan untuk menemukan pola akan dilakukan berikut ini langkah-langkah yang dilakukan :

- 1) *Memanggil pandas dan MLxtend, setelah itu panggil dataset untuk dibaca*

```
import pandas as pd
from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules
df = pd.read_excel('http://archive.ics.uci.edu/ml/machine-learning-databa
ses/00352/Online%20Retail.xlsx')
df.head()
```

2) *Melakukan data cleaning*

```
df['Description'] = df['Description'].str.strip()
df.dropna(axis=0, subset=['InvoiceNo'], inplace=True)
df['InvoiceNo'] = df['InvoiceNo'].astype('str')
df = df[~df['InvoiceNo'].str.contains('C')]
```

3) *Setelah pembersihan, kita perlu mengkonsolidasikan item menjadi 1 transaksi per baris dengan masing- masing produk 1 dikodekan.*

```
basket = (df[df['Country'] == "France"]
          .groupby(['InvoiceNo', 'Description'])['Quantity']
          .sum().unstack().reset_index().fillna(0)
          .set_index('InvoiceNo'))
```

4) *Dikarenakan banyak sekali angka 0 didalam data tetapi kami juga perlu memastikan bahwa setiap nilai positif dikonversi menjadi 1 dan kurang dari 0 diatur ke 0. Maka kode dibawah ini dijalankan.*

```
def encode_units(x):
    if x <= 0:
        return 0
    if x >= 1:
        return 1
basket_sets = basket.applymap(encode_units)
basket_sets.drop('POSTAGE', inplace=True, axis=1)
```

5) *Setelah data terstruktur dengan benar, maka rule siap diterapkan kedalam data tersebut.*

```
rules = association_rules(frequent_itemsets, metric="lift", min_thresh
old=1)
rules
```

6) *Setelah rule atau pola telah terhitung kita dapat melakukan filter pada data tersebut.*

```
rules[ (rules['lift'] >= 6) &
       (rules['confidence'] >= 0.8) ]
```

Untuk penerapan metode hotelling kedalam itemset masih belum dilakukan karena keterbatasan waktu, dan yang hanya dapat dilakukan hanya mencari pola atau itemset dari data penjualan tersebut.

II. HASIL DAN PEMBAHASAN

Dari penerapan metode market basket analysis, ditemukan dua pola yang menarik yaitu yang pertama jam alarm berwarna merah dan hijau, sedangkan pola kedua yaitu, gelas kertas berwarna merah, serbet dan piring. Hasil dari penerapan metode market basket analysis dapat dilihat dibawah ini.

| | antecedants | consequents | support | confidence | lift |
|----|---|--------------------------------------|----------|------------|----------|
| 4 | (SET/6 RED SPOTTY PAPER PLATES) | (SET/6 RED SPOTTY PAPER CUPS) | 0.127551 | 0.960000 | 6.968889 |
| 5 | (SET/6 RED SPOTTY PAPER CUPS) | (SET/6 RED SPOTTY PAPER PLATES) | 0.137755 | 0.888889 | 6.968889 |
| 6 | (SET/6 RED SPOTTY PAPER CUPS, SET/20 RED RETRO... | (SET/6 RED SPOTTY PAPER PLATES) | 0.102041 | 0.975000 | 7.644000 |
| 7 | (SET/6 RED SPOTTY PAPER PLATES, SET/6 RED SPOT... | (SET/20 RED RETROSPOT PAPER NAPKINS) | 0.122449 | 0.812500 | 6.125000 |
| 8 | (SET/6 RED SPOTTY PAPER PLATES, SET/20 RED RET... | (SET/6 RED SPOTTY PAPER CUPS) | 0.102041 | 0.975000 | 7.077778 |
| 18 | (ALARM CLOCK BAKELIKE GREEN) | (ALARM CLOCK BAKELIKE RED) | 0.096939 | 0.815789 | 8.642959 |
| 19 | (ALARM CLOCK BAKELIKE RED) | (ALARM CLOCK BAKELIKE GREEN) | 0.094388 | 0.837838 | 8.642959 |
| 24 | (SET/6 RED SPOTTY PAPER PLATES) | (SET/20 RED RETROSPOT PAPER NAPKINS) | 0.127551 | 0.800000 | 6.030789 |

Gambar III-1

Pola-pola yang telah didapatkan tadi nantinya akan disusun agar saling berdekatan sehingga pembeli akan tertarik untuk membeli itemset yang telah disusun tersebut. Atau dapat penjual dapat menerapkan promo atau discount terhadap itemset tersebut sehingga pelanggan akan tertarik membeli salah satu item yang nantinya akan mendorong pembeli untuk membeli dua item secara sekaligus.

Dikarenakan keterbatasan waktu penelitian, tidak bisa untuk menerapkan hasil dari metode market basket analysis dan juga metode hotelling kedalam itemset yang telah ditemukan. Sehingga untuk mengetahui dampak dari kedua metode tersebut masih belum diketahui.

III. KESIMPULAN

Penelitian ini telah menghasilkan pola atau itemset dari data dummy yang telah digunakan. Dari metode market basket analysis ditemukan dua pola, pola yang pertama jam alarm berwarna merah dan hijau, sedangkan pola kedua yaitu, gelas kertas berwarna merah, serbet dan piring. Akan tetapi keberhasilan dari penelitian ini masih belum bisa ditentukan karena, belum diterapkannya metode hotelling dan juga penghitungan peningkatan pembelian masih belum diketahui.

Untuk penelitian selanjutnya diharapkan data yang digunakan adalah data transaksi ikan, selain itu pemilihan tempat penjualan atau pengambilan data transaksi harus diperhatikan karena untuk penerapan metode hotelling jarak antara dua tempat pengambilan data akan sangat mempengaruhi hasil dari penelitian. Dan diharapkan untuk penelitian-penelitian selanjutnya akan menyempurnakan

metode dan berhasil dalam melakukan penelitian ini. Dampak dari penelitian ini akan sangat besar dikarenakan jika pemerintah melakukan revitalisasi pada pasar-pasar tradisional metode ini dapat diterapkan kedalam pasar tersebut.

IV. DAFTAR PUSTAKA

1. E. Turban, *Decision Support Systems and Intelligent Systems*, New Jersey: Prentice Hall, 2001, p. 867.
2. J. Han, M. Kamber and J. Pei, *Data Mining: Concepts and Techniques*, Massachusetts: Morgan Kaufmann, 2011.
3. D. T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining*, New York: Wiley-Interscience, 2004.
4. M. Bramer, *Principles of Data Mining*, London: Springer-Verlag, 2007.