

## BAB 2

### Tinjauan Pustaka

#### 2.1 Penelitian Terdahulu

Tabel 2.1 merupakan daftar penelitian terdahulu yang pernah dilakukan oleh peneliti lain terkait dengan penerapan *data mining clustering* untuk membentuk sebuah segmentasi pelanggan yang dihasilkan dari beberapa data yang masih berhubungan dengan karakteristik pelanggan.

Tabel 2.1 Penelitian Terdahulu

No	Judul Penelitian	Ringkasan	Hasil	Kaitan Penelitian
1.	<i>Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix</i> (2015)	Penelitian ini menjelaskan bagaimana ke – 5 obyek <i>marketing mix</i> dapat digunakan sebagai dasar analisa <i>big data</i> untuk <i>business intelligence</i>	Usualan <i>framework</i> untuk memandu peneliti yang memiliki fokus pada <i>big data analytics for Business Intelligence</i>	Penggunaan <i>marketing mix framework for big data management</i> pada Gambar 1.1 sebagai panduan dasar dalam pemilihan data, metode, dan pengaplikasian
2	<i>From “foodies” to “cherry-pickers”: A clustered-based segmentation of specialty food retail customers.</i> (2018)	Penelitian ini memberikan pendapat bahwa kategorisasi pelanggan ritel yang beragam dapat membantu para penyedia pemilik usaha dalam menentukan pangsa pasar yang lebih baik, metode yang dipakai dalam penelitian ini adalah klasterisasi menggunakan K – Means Algorithm	Menghasilkan 4 <i>cluster</i> konsumen potensial	Penggunaan teknik klaster dengan algoritma K - Means
3	<i>Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis</i> (2016)	Penelitian ini memberikan pendapat bahwa RFM <i>analysis</i> dapat membantu pihak eksekutif hotel untuk meningkatkan strategi pemasaran mereka dalam mencapai target perusahaan	Pengetahuan tentang karakter pelanggan hotel potensial ke dalam 8 kelompok klaster yang didefinisikan menjadi beberapa klasifikasi penamaan pelanggan	Penggunaan RFM analisis ke dalam proses pengolahan data
4	<i>RFM ranking – An effective approach to customer segmentation</i> (2018)	Penelitian ini melakukan perbandingan data hasil analisa RFM ke dalam 3 (tiga) algoritma klaster, yaitu K – Means, Fuzzy C – Means, dan RF K – Means	bahwa algoritma K – Means merupakan algoritma yang diusulkan, karena dapat memberikan waktu eksekusi yang lebih sedikit dan juga mengurangi jumlah iterasi	Pemilihan algoritma K - Means yang disarankan, karena ada beberapa kelebihan dibandingkan dengan beberapa algoritma lainnya
5	<i>An approach to optimised customer segmentation and profiling using RFM, LTV, and demographic features</i> (2012)	Berdasarkan hasil kajian pustaka terdahulu, melalui penelitian ini peneliti mengusulkan 3 <i>framework</i> untuk penggolongan pelanggan berdasarkan kombinasi beberapa analisa pengolahan data yang bersumber dari data transaksi dan data profil pelanggan. Metode yang dipakai adalah CRM, RFM, dan LTV	Hasil komparasi dari ke 3 <i>framework</i> tersebut, menghasilkan bahwa ada 1 <i>framework</i> yang disarankan, yaitu hasil analisa RFM kemudian dikelompokkan kedalam masing – masing segmen RFM menggunakan tambahan data demografi pelanggan	Penggunaan data transaksi yang kemudian dikombinasikan dengan data demografi pelanggan

## 2.2 Landasan Teori

### 2.2.1 *Bisnis On Demand Service*

*On demand service* merupakan model bisnis dari sebuah perusahaan untuk memenuhi permintaan konsumen melalui penyediaan langsung barang dan jasa sesuai permintaan (Kielstra, 2016). *Bisnis on demand service* yang didasarkan atas kesesuaian dengan permintaan pelanggan atau konsumen yang dihimpun pada satu aplikasi berbasis *online* yang diharapkan dapat membantu memudahkan akses pelanggan (Kielstra, 2016).

Adapun manfaat utama yang ditawarkan melalui model bisnis ini yaitu, memberikan pelanggan peluang besar untuk melakukan pemesanan dalam waktu yang tergolong cepat, murah, nyaman, dan transparan. Alur prosesnya pun juga dirancang untuk memudahkan dalam pemakaian, mulai dari mencari, memesan, menjadwalkan, membayar dan memberikan ulasan terhadap hasil layanan, yang kesemua hal ini dapat ditangani dalam satu aplikasi (Inc., 2019).

### 2.2.2 Segmentasi Pelanggan

Segmentasi pelanggan atau juga bias disebut segmentasi pasar merupakan suatu cara atau pola untuk menghimpun pelanggan dengan mengelompokkan pelanggan yang memiliki karakteristik serupa (Fisk, 2015). Terdapat dua faktor yang dapat digunakan untuk melakukan pengelompokan pelanggan, antara lain :

1. Kebutuhan (*needs*): pengelompokan yang didasarkan pada apa yang pelanggan butuhkan sehingga diharapkan dapat memuaskan mereka.
2. Karakteristik pembelian (*buying characteristics*): mengetahui karakter pelanggan pada saat melakukan transaksi seperti *messaging*, *marketing channels*, dan *sales channels* untuk dijadikan pendekatan sebagai masukan untuk menjual kepada pelanggan secara kompetitif, tepat sasaran dan ekonomis (Herbert, n.d.).

Ada empat poin keberhasilan untuk melakukan segmentasi pelanggan antara lain :

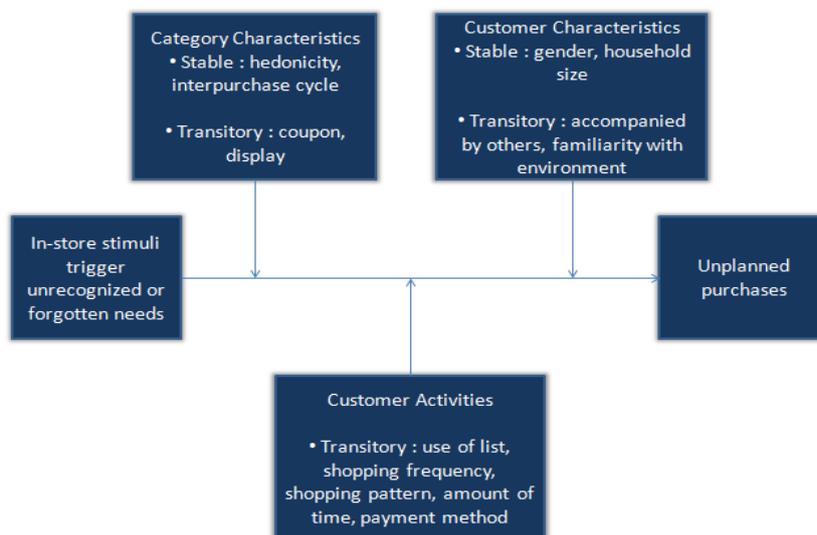
1. Memahami biaya bisnis : mengetahui bukan hanya tentang biaya bisnis secara keseluruhan, tetapi juga biaya yang setiap detail transaksi pelanggan. Hal ini diharapkan membantu meningkatkan *customers profit potential* dan *customer lifetime values* (CLV)

2. Identifikasi pelanggan terbaik : mengetahui pelanggan yang potensial dan memberikan pelayanan serta rekomendasi terbaik yang diharapkan dapat memberi nilai kepuasan pada pelanggan
3. Mengumpulkan data dari sumber eksternal: mengumpulkan setiap informasi terkait transaksi pelanggan untuk menciptakan pengalaman berbelanja yang relevan dan tepat sasaran
4. Membuat profil pelanggan: ini ditujukan untuk memahami kebutuhan, tantangan dan pola pikir terhadap pelanggan (Fisk, 2015).

### 2.2.3 Karakteristik Pelanggan

Karakteristik pelanggan adalah pola sifat menguntungkan dari pelanggan yang berupa informasi yang dapat digunakan untuk membantu pengembangan produk atau layanan baru dengan mengacu pola yang didapat dari pelanggan (Schuurman, Mahr, & De Marez, 2011).

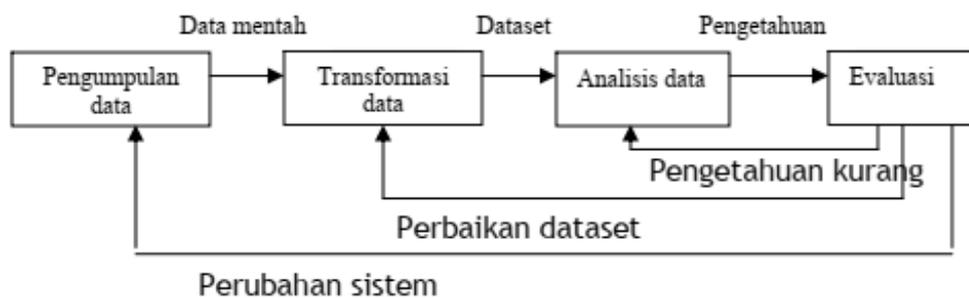
Karakteristik dapat dicontohkan seperti status, kualifikasi, spesialisasi, sifat pekerjaan, kegiatan dan prestasi profesional dari pelanggan (Suhadi & Wandebori, 2012). Karakteristik pelanggan dapat digunakan untuk menentukan target pasar lebih spesifik dan dapat membuat strategi pemasaran yang cocok untuk pelanggan mereka (Sridhar, 1988), serta karakteristik dapat menguntungkan dan mengoptimalkan keterlibatan pengguna dalam proses inovasi pelayanan pelanggan (Inman, Winer, & Ferraro, 2009). Pada Gambar 2.2.1 dapat dilihat pola informasi pelanggan yang dapat dijadikan rujukan untuk membangun karakteristik yang dibutuhkan sebagai peningkatan pelayanan terhadap pelanggan.



Gambar 2.2.1 Bagan karakteristik pelanggan (Inman et al., 2009)

#### 2.2.4 Data Mining

*Data mining* adalah bagian dari *knowledge discovery in databases* (KDD) yang terdiri dari beberapa proses seperti *data cleaning*, *data selection*, *data integration*, *data transformation*, *data mining*, *pattern evaluation*) dan *knowledge presentation* (Firdaus, 2017). *Data mining* dapat digunakan sebagai salah satu cara praktis untuk automasi data dari berbagai *source* seperti *database*, *data warehouse* dan penyimpanan data lainnya (Han & Kamber, n.d.). Adapun langkah-langkah proses pada *data mining* dapat dilihat pada Gambar 2.2.2.

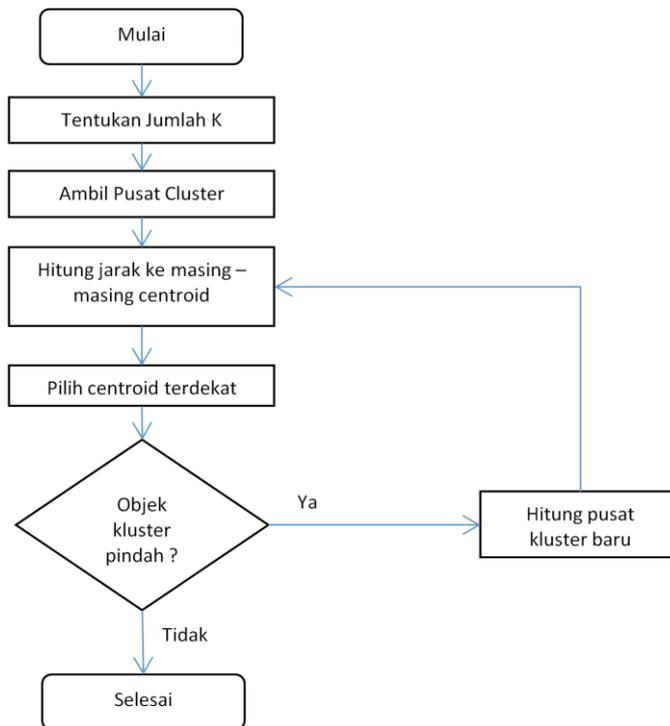


Gambar 2.2.2 Proses dari data mining (Ayub, 2007)

Pada *data mining*, pola atau informasi yang dibutuhkan untuk proses pengembangan pelayanan dapat dipilih dengan menggunakan teknik atau metode tertentu (algoritma) (Muzakir & Wulandari, 2016). Analisa data informasi yang dikumpulkan diharapkan dapat mencapai tujuan bisnis untuk meningkatkan kepuasan pelanggan sesuai dengan target dan tepat sasaran (Firdaus, 2017).

#### 2.2.5 K – Means Algorithm

*K-Means Algorithm* merupakan salah satu algoritma pembelajaran sederhana yang prosedurnya efisien dan mudah (Kolay & Ray, 2017). Pada Gambar 2.2.3 dapat dilihat diagram alur yang menggambarkan bagaimana algoritma K-Means itu berjalan dari awal hingga akhir.



Gambar 2.2.3 Bagan algoritma K-Means (Raghuwanshi & Arya, 2012)

Algoritma K-Means melakukan pengelompokan data sesuai dengan kemiripan data dalam kelompok yang sama dan ketidakmiripan dengan kelompok data yang berbeda (Larose, 2004), algoritma K-Means melakukan klasifikasi data dengan cara pengelompokan data apriori tetap untuk mendapatkan hasil yang berbeda (Raghuwanshi & Arya, 2012). Adapun langkah-langkah pada algoritma K-Means antara lain (Ghosh & Dubey, 2013) :

1. Set K: memilih sejumlah data untuk diklaster.
2. Inisialisasi : memilih K titik awal yang digunakan sebagai perkiraan awal cluster centroid.
3. Tetapkan K pada pusat dataset yang akan diklaster.
4. Perhitungan centroid: setiap titik dalam dataset ditugaskan ke sebuah cluster, perlu untuk menghitung ulang k yang baru.
5. Langkah 3 dan 4 perlu diulang sampai tidak ada titik yang mengubah klasternya atau sampai centroid tidak lagi bergerak.

### 2.2.6 Penentuan K optimal dengan Elbow

Salah satu cara dalam menentukan jumlah k dalam suatu *dataset* adalah dengan menggunakan metode Elbow, metode Elbow merupakan perbandingan antara jumlah

klaster yang akan membentuk siku pada suatu titik, jika penurunan dan membentuk siku paling besar, maka nilai klaster tersebut yang terbaik (Tibshirani, 2001).

### 2.2.7 Mengukur jarak koordinat dengan rumus Haversine

Haversine merupakan sebuah rumus persamaan dalam bidang navigasi untuk menghitung jarak busur antara dua titik dari garis bujur (*longitude*) dan garis lintang (*latitude*) dimana hasil dari perhitungan jarak tersebut adalah jarak terpendek antara kedua titik tersebut (Yulianto, Ramadiani, & Kridalaksana, 2018). Persamaan (1) merupakan rumus perhitungan menggunakan Haversine.

$$\begin{aligned} a &= \sin^2(\Delta\phi/2) + \cos \phi_1 \cdot \cos \phi_2 \cdot \sin^2(\Delta\lambda/2) \\ c &= 2 \cdot a \tan^{-1} \left( \sqrt{a}, \sqrt{1-a} \right) \\ d &= R \cdot c \end{aligned} \quad (1)$$

Keterangan :

$\phi$  = *latitude*,

$\lambda$  = *longitude*,

R = *radius bumi* (*radius* = 6,371 km)

### 2.2.8 Normalisasi data dengan *Min Max Normalization*

Standarisasi data untuk semua variabel yang digunakan dalam penelitian ini menggunakan metode *Min Max Normalization*. *Min Max Normalization* ini akan mempermudah perbandingan nilai yang memiliki satuan ukuran yang berbeda (Virmani, 2015), dimana varian variabel yang ada pada penelitian ini menggunakan satuan ukuran yang berbeda seperti pada Tabel 4.2. Persamaan (2) merupakan rumus Normalisasi Min – Max.

$$z = \frac{x - \min(x)}{[\max(x) - \min(x)]} \quad (2)$$

Keterangan :

z = hasil normalisasi,

x = nilai x (asli),

min(x) = nilai minimal untuk variabel x,

max(x) = nilai maksimal untuk variabel x.