

Analisis Pemesanan *Hosting* PT Qwords Company International Menggunakan Metode *K-Means Clustering*

Cahyo Prianto¹, Shinta Amelia²

^{1,2} Politeknik Pos Indonesia

¹ cahyoprianto@poltekpos.ac.id, ² shinta.amelia216@gmail.com

ABSTRAK

Qwords.com adalah perusahaan penyedia jasa layanan *Web Hosting* yang melayani registrasi nama domain dan *web hosting*. Qwords.com berdiri pada 10 Juli 2005, dan mulai menggunakan *brand* Qwords sejak 28 Agustus 2005. *Hosting* adalah tempat untuk membuat halaman *website* yang telah dibuat *online*, sehingga bisa diakses oleh orang lain. Dalam penelitian ini, penulis menemukan masalah, yaitu sulitnya menemukan *client* yang sering memesan *hosting* dan dari daerah mana *client* itu berasal. Sehingga perusahaan sulit untuk menentukan target promosi yang akan dilakukan. Oleh karena itu, diperlukan suatu metode yang dapat memecahkan masalah yang ada di perusahaan. Dalam penelitian ini, penulis menggunakan metode *K-Means Clustering* untuk memecahkan masalah yang ada. Dimana metode ini berguna untuk memberikan *output* berupa hasil perhitungan yang menunjukkan produk yang paling banyak dibeli dalam kategori paket *hosting* tahun 2017, dan nantinya data ini berguna untuk menentukan strategi promosi.

Kata Kunci: Analisis, *Hosting*, *K-Means Clustering*

ABSTRACT

Qwords.com is a web hosting service company that serves domain name registration and web hosting. Qwords.com was established on July 10, 2005, and started using the Qwords.com brand since August 28, 2005. Hosting is a place to create web pages that have been made online, so that others can access them. In this study, the author found a problem, namely the difficulty of finding clients who often host in the area and from the company where the clients from. So the company is difficult to determine the target of promotion that will be done. Therefore, we need a method that can solve problems that exist to provide output in the form of calculation results that show the most purchased products in the category of hosting packages in 2017, and later this data is useful to determine the target of promotion.

Keywords: Analysis, *Hosting*, *K-Means Clustering*.

1. PENDAHULUAN

Promosi adalah salah satu hal yang perlu dilakukan oleh suatu perusahaan untuk meningkatkan penjualan. Dalam melakukan promosi, perusahaan harus mengetahui produk apa yang penjualannya lebih banyak dan produk apa yang penjualannya paling sedikit. Banyaknya data akan mempersulit perusahaan untuk mengetahui produk apa yang paling banyak dipesan, dan produk mana yang peminatnya sedikit.

Qwords.com adalah perusahaan penyedia jasa layanan *web hosting* yang melayani registrasi nama *domain* dan *web hosting* sejak 28 Agustus 2005. *Hosting* adalah salah satu bentuk layanan jasa penyewaan tempat di internet yang memungkinkan perorangan atau organisasi menampilkan layanan jasa atau produknya di internet. Dengan kata lain, *hosting* adalah sebuah wadah atau tempat sebuah *website* [1].

Pemesanan *hosting* pada Qwords.com terus meningkat dan pertumbuhan data juga akan terus meningkat dan akan menimbulkan penumpukan data dalam *database*. Dengan menumpuknya data pada *database*, akan mengakibatkan perusahaan kesulitan untuk menemukan produk yang paling banyak

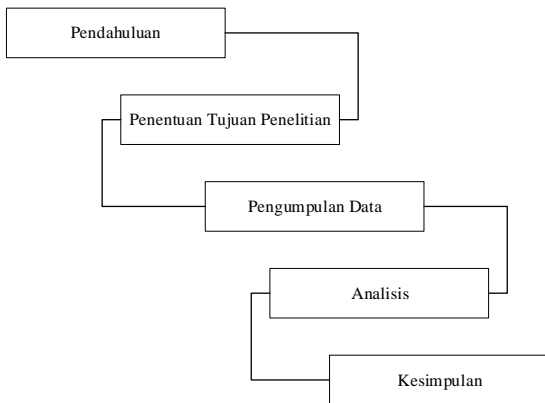
diminati dan dari provinsi mana pemesannya. Oleh sebab itu, data mining dapat diterapkan untuk menganalisa data untuk nantinya akan dilakukan strategi promosi yang cocok untuk Qwords.com.

Untuk dapat mengetahui provinsi mana yang paling banyak memesan paket *hosting*, maka akan dilakukan penggalan data sehingga dapat diperoleh informasi yang berguna untuk proses analisis. Penelitian ini menggunakan metode *K-Means Clustering*, dimana data akan dibagi menjadi beberapa *cluster*. Dengan adanya metode ini, maka akan mudah untuk menentukan strategi promosi yang tepat untuk mempromosikan produk *hosting* dari Qwords.com.

Penelitian ini hanya meneliti tentang kategori produk *hosting*, mencakup analisis dan pengolahan data. Analisis dan pengolahan data pada penelitian ini menggunakan dua cara, yang pertama yaitu menggunakan aplikasi pengolah data WEKA, dan menggunakan cara manual yaitu dengan menggunakan perhitungan *excel*. Hasil dari perhitungan ini nantinya akan digunakan untuk menentukan strategi promosi yang tepat untuk produk Qwords.com.

2. METODOLOGI PENELITIAN

Metodologi penelitian adalah teknik dalam penelitian dan digunakan dalam pemecahan masalah. Pada penelitian ini metodologi penelitian digunakan untuk memecahkan masalah yang ditemukan [2]. Metode yang digunakan dalam penelitian ini adalah metode pengumpulan data dan informasi mengenai data *hosting* yang telah dibayar pada tahun 2017.



Gambar 1. Diagram Alur Metodologi Penelitian

Setelah data dan informasi didapatkan, maka data tersebut akan diolah menggunakan metode *K-Means Clustering*, dimana metode ini akan mengelompokkan suatu data kedalam *cluster* tertentu. *Clustering* digunakan untuk mengidentifikasi kelompok data dari sebuah kasus yang didasarkan pada sebuah kelompok atribut, dan mengelompokkan data yang memiliki kemiripan atribut [3].

Langkah-langkah penerapan metode *K-Means clustering* yaitu :

1. Menentukan jumlah *cluster*.
2. Menentukan titik pusat awal di setiap *cluster*, yang ditentukan secara *random*.
3. Hitung jarak dari setiap data. Disini akan diketahui *cluster* mana yang paling dekat dengan data.

Berikut rumus perhitungannya

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Keterangan :

D(i,j) = Jarak data ke i ke pusat *cluster* j

X_{ii} = Data ke i pada atribut data ke k

X_{ij} = Titik pusat ke j pada atribut ke k

4. Menghitung kembali kriteria dengan titik pusat *cluster*.

5. Hasil perhitungan akan berhenti jika antara iterasi sebelumnya tidak ada lagi perpindahan data.

3. HASIL DAN PEMBAHASAN

Tujuan penelitian ini adalah untuk menentukan produk yang paling banyak dipesan, dan juga mengetahui provinsi yang paling banyak memesan produk *hosting*, yang hasilnya digunakan sebagai dasar dalam menentukan strategi promosi yang tepat untuk Qwords.com.

Penerapan Algoritma *K-Means Clustering*

Dalam penelitian ini, data yang diolah terdapat 100 data *hosting* yang telah dibayar pada tahun 2017. Penerapan metode ini menggunakan dua atribut yaitu provinsi dan total pemesanan. Data tersebut dapat dilihat pada tabel 1 :

Tabel 1. Data mentah

No	Packageid	Kota	Provinsi	Total
1	168	Bandung	Jawa Barat	699
2	169	Jakarta	DKI Jakarta	97
3	170	Bandung	Jawa Barat	54
4	171	Kelapa Dua	Banten	148
5	172	Sleman	DI Yogyakarta	238
6	173	Cimahi	Jawa Barat	217
7	174	Aceh	Banda Aceh	38
8	177	Jakarta	DKI Jakarta	41
9	274	Timur	DKI Jakarta	42
10	329	Jakarta Pusat	DKI Jakarta	42
		Bandung	Jawa Barat	350

Untuk kolom kedua, terdapat *packageid*. *Packageid* adalah nama paket *hosting* yang disediakan oleh Qwords.com. Nama paket-paket tersebut dibedakan berdasarkan kapasitas yang dimiliki dari masing-masing paket *hosting*. Daftar nama paket *hosting* yang diolah disajikan pada tabel 2 :

Tabel 2. *Packageid*

<i>Packageid</i>	Nama Paket	Kapasitas
168	Cloud Web <i>Hosting</i> L	15GB
169	Cloud Web <i>Hosting</i> XL	30GB
170	Cloud Web <i>Hosting</i> XXL	40GB
171	Cloud Web <i>Hosting</i> XS	500MB
172	Cloud Web <i>Hosting</i> S	1.5GB
173	Cloud Web <i>Hosting</i> M	3GB
174	Cloud Web <i>Hosting</i> L USA	15GB
177	Cloud Web <i>Hosting</i> L Singapura	15GB
274	Cloud Web <i>Hosting</i> Base WP	2GB
	<i>Hosting</i>	
329	Cloud Web <i>Hosting</i> E	1GB

Inisialisasi Data

Pada penelitian ini, atribut yang digunakan adalah provinsi dan total. Dalam *K-Means Clustering*, data yang bisa diolah hanya bisa

menggunakan data numerik. Agar data tersebut bisa diolah, maka dilakukan inisialisasi data. Qwords.com adalah perusahaan yang berkantor pusat di wilayah Bandung, Jawa Barat. Oleh karena itu, untuk inisialisasi provinsi, kota yang berada pada wilayah Jawa Barat akan diberikan inisial 100, yang terdekat kedua diberikan inisialisasi 90, begitu seterusnya hingga provinsi dengan jarak paling jauh dari Jawa Barat. Hasil dari inisialisasi yang dilakukan, disajikan pada tabel 3:

Tabel 3. Inisialisasi Provinsi

Provinsi	Inisialisasi
Jawa Barat	100
DKI Jakarta	90
Banten	80
DI Yogyakarta	70
Jawa Tengah	60
Jawa Timur	50
Sumatera Selatan	40
Nusa Tenggara Barat	30
Banda Aceh	20
Sulawesi Selatan	10

Setelah dilakukan inisialisasi data, maka data tersebut dapat diolah menggunakan metode *K-Means Clustering* [3].

1. Menentukan jumlah *cluster*. Dalam penelitian ini, data dikelompokkan menjadi tiga *cluster*.
2. Menentukan titik pusat awal di setiap *cluster*, yang ditentukan secara random, seperti pada tabel 4.

Tabel 4. Titik Pusat Awal

	Total	Provinsi
C1	349	100
C2	54	100
C3	493	90

Tabel 5. Hasil Algoritma *K-Means Clustering*

Pid	Provinsi	Jumlah	Inisialisasi	DC1	DC2	DC3	CLUSTER
168	Jawa Barat	699	100	446	643	169	C3
169	DKI Jakarta	97	90	156	41	434	C2
170	Jawa Barat	54	100	199	12	477	C2
171	Banten	148	80	64	150	332	C2
172	DI Yogyakarta	238	70	23	183	293	C1
173	Jawa Barat	217	100	38	161	315	C1
174	Banda Aceh	38	20	225	70	497	C2
177	DKI Jakarta	41	90	212	15	490	C2
274	DKI Jakarta	42	90	211	14	489	C2
329	Jawa Barat	350	100	83	278	182	C1

Dengan pusat *cluster* pada iterasi ketiga disajikan dalam Tabel 6 :

3. Hitung jarak dari setiap data. Disini akan diketahui *cluster* mana yang paling dekat dengan data.

$$D(1,1) = \sqrt{(699 - 253)^2 + (100 - 88)^2} = 446$$

Dari hasil perhitungan diatas didapatkan bahwa jarak data *hosting* yang pertama dengan pusat *cluster* yang pertama adalah 446.

$$D(1,2) = \sqrt{(699 - 56)^2 + (100 - 88)^2} = 643$$

Dari hasil perhitungan diatas diperoleh bahwa jarak data *hosting* yang pertama dengan pusat *cluster* yang kedua adalah 643.

$$D(1,3) = \sqrt{(699 - 531)^2 + (100 - 82)^2} = 169$$

Dari hasil perhitungan diatas didapatkan bahwa jarak data *hosting* yang pertama dengan pusat *cluster* yang kedua adalah 169.

4. Menghitung kembali kriteria dengan titik pusat *cluster*, tetapi untuk perhitungan ini nilai dari C1,C2, dan C3 akan berubah sesuai dengan hasil dari perhitungan yang pertama.
5. Hasil perhitungan akan berhenti jika antara iterasi terakhir dan iterasi sebelumnya tidak ada lagi perpindahan data. Pada perhitungan ini, iterasi berhenti pada iterasi ketiga.

Pengolahan Data Menggunakan Cara Manual (Excel)

Tabel 5 menunjukkan perhitungan lengkap *K-Means Clustering* :

Tabel 6. Titik Pusat Akhir

	Total	Provinsi
C1	268	88
C2	72	86
C3	531	81

Berdasarkan hasil pengelompokan data menggunakan metode *K-Means Clustering*,

Tabel 7. Penggolongan *Cluster*

<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>
Jawa Barat = 22	Jawa Barat = 20	Jawa Barat = 5
DKI Jakarta = 5	DKI Jakarta = 18	DKI Jakarta = 3
Jawa Tengah = 2	Jawa Tengah = 1	DI Yogyakarta = 1
Banten = 1	Banten = 4	Sumatera Selatan = 1
Jawa Timur = 3	Jawa Timur = 2	Sulawesi Selatan = 1
DI Yogyakarta = 2	DI Yogyakarta = 5	Banten = 1
Nusa Tenggara Barat = 1	Sulawesi Selatan = 1	
	Banda Aceh = 1	

Dari hasil *clustering* yang didapatkan adalah, pada *cluster 1* didominasi oleh Provinsi Jawa Barat dengan jumlah 22 kali transaksi, disusul dengan DKI Jakarta sebanyak 5 kali transaksi, kemudian Jawa Tengah dengan 2 kali transaksi.

Sedangkan untuk *cluster* kedua didominasi oleh Provinsi Jawa Barat dengan 20 kali transaksi, disusul dengan DKI Jakarta sebanyak 18 kali transaksi, kemudian DI Yogyakarta sebanyak 5 kali transaksi.

Kemudian untuk *cluster* ketiga didominasi oleh Provinsi Jawa Barat dengan jumlah transaksi sebanyak 5 kali transaksi, DKI Jakarta sebanyak 3 kali transaksi, dan DI Yogyakarta sebanyak 1 kali transaksi.

Pengolahan Data Menggunakan Aplikasi Pengolah Data (WEKA)

Weka merupakan singkatan dari *Waikato Environment of Knowledge Analysis*. *Weka* adalah aplikasi data mining open source berbasis java. *Weka* terdiri dari koleksi algoritma *machine learning* yang dapat digunakan untuk melakukan generalisasi atau formulasi dari sekumpulan data *sampling* [4].

Berikut hasil perhitungan menggunakan WEKA yang disajikan dalam gambar 2.

```
Number of iterations: 3
Within cluster sum of squared errors: 5.2232121894646655

Initial starting points (random):

Cluster 0: 349,100
Cluster 1: 54,100
Cluster 2: 493,90

Missing values globally replaced with mean/mode
```

didapatkan hasil *clustering* hingga iterasi ketiga, dimana titik pusat sudah berubah dan tidak ada lagi data yang berpindah antar *cluster*, seperti yang terlihat pada tabel 7.

```
Final cluster centroids:
Attribute      Full Data      Cluster#
                (100.0)      (35.0)      (53.0)      (12.0)
-----
jumlah         184.09      257.7429      58.9434      522
Inisialisasi   86.3        91.1429      85.4717      75.8333

Time taken to build model (full training data) : 0.01 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      35 ( 35%)
1      53 ( 53%)
2      12 ( 12%)
```

Gambar 2. Hasil *Clustering* data *hosting* menggunakan *K-Means Clustering*

- Keterangan :
1. *Number of Iteration* : Jumlah Iterasi
 2. *Within Cluster sum of squared errors = SSE*, yaitu jumlah kuadrat *error* yang dihasilkan, semakin kecil kuadrat *error* yang dihasilkan, maka data akan semakin *valid*.
 3. Untuk hasil yang dihasilkan, pada *cluster 0* adalah sebanyak 35 kali transaksi, pada *cluster 1* sebanyak 53 kali transaksi, kemudian pada *cluster 2* sebanyak 12 kali transaksi.

Hasil akhir kelompok *cluster* yang diperoleh antara proses perhitungan menggunakan ms.excel adalah sama dengan proses perhitungan menggunakan aplikasi *weka*.

4. KESIMPULAN DAN SARAN
 Metode *K-Means Clustering* adalah metode yang tepat untuk mengelompokkan data paket

hosting, agar perusahaan dapat mengelompokkan produk dan provinsi mana yang paling sering melakukan transaksi, sehingga memudahkan perusahaan untuk menentukan strategi promosi.

Informasi yang diperoleh dari hasil proses *clustering* tersebut adalah pada data pemesanan produk yang semula ada 100 *record* data dikelompokkan menjadi tiga *cluster*.

1. *Cluster* pertama terdiri dari 36 *record*, transaksi, dengan *packageid* 329,172,dan 173.
2. *Cluster* kedua terdiri dari 52 *record* transaksi dengan *packageid* 171,169,274,337,177,dan 170.
3. *Cluster* ketiga terdiri dari 12 *record* transaksi dengan *packageid* 168.

Strategi promosi yang bisa diterapkan yaitu :

1. Melakukan promosi ke kota yang berada pada provinsi yang berada pada *cluster* kedua, yang didominasi oleh provinsi Jawa Barat, DKI Jakarta, dan DI Yogyakarta. Dengan melakukan promosi tersebut, perusahaan bisa meningkatkan laba dari penjualan produk *hosting*.
2. Pihak *marketing* bisa lebih fokus untuk mempromosikan produknya ke kota-kota yang paling banyak memesan produk *hosting*, sehingga promosi yang dilakukan akan lebih efektif dan efisien karena dengan melakukan strategi ini pihak *marketing* dapat membagi tim untuk melakukan promosi ke kota-kota yang paling banyak memesan produk *hosting*.

5. REFERENSI

- [1] F. Aditya and E. Handoyo, "Perbandingan Paid Hosting Dan Free Hosting Berdasarkan Fasilitas Backup yang Ada," *J. Teknol.*, vol. 4, no. 1, pp. 1–5, 2011.
- [2] E. Utami, Suwanto Raharjo, and Jazi Eko Istiyanto, "Metodologi Penelitian Pada Ilmu komputer," in *Seminar Nasional Teknologi 2007*, 2007.
- [3] J. O. Ong, "Implementasi Algoritma K-means clustering untuk menentukan strategi marketing president university," *J. Ilm. Tek. Ind.*, vol. 12, no. 01, pp. 10–20, 2013.
- [4] S. Adinugroho and Yuita Arum Sari, *Implementasi Data Mining Menggunakan Weka*. Malang: Universitas Brawijaya Press, 2018.