

## Prediksi Kebutuhan Buah dengan Segmentasi Pasar Menggunakan K-Means

Jejen Arisandi<sup>1</sup>

Magister Teknik Informatika  
Fakultas Ilmu Komputer, Universitas Sriwijaya  
Palembang, Indonesia  
JejenArisandi@gmail.com

Samsuryadi<sup>2</sup>

Teknik Informatika Fasilkom, Universitas Sriwijaya  
Jl. Raya Palembang – Prabumulih Km.32 Indralaya, OI line  
Palembang, Indonesia  
syamsuryadi@unsri.ac.id

**Abstrak—** *K-Means diterapkan untuk melakukan prediksi segmentasi pasar berdasarkan parameter masukan untuk menentukan kelompok, yang ditinjau berdasarkan derajat similaritas antar elemen dalam satu kelompok tinggi sedangkan tingkat similaritas anggota pada kelompok lain rendah. Seleksi attribute dilakukan secara berbeda-beda agar hasil yang didapat menjadi lebih akurat. Similaritas elemen terhadap kelompok ditentukan dengan kedekatan objek terhadap nilai rata-rata pada kelompok atau pusat massa. Algoritma K-means dapat diterapkan untuk memprediksi segmentasi pasar berdasarkan faktor kemasan, promosi produk, waktu penyebaran, identifikasi pelanggan yang potensial dan jumlah penjualan yang telah dilakukan. Hasil awal segmentasi pasar telah terbentuk 3 kelompok dengan persentasi yang tidak jauh berbeda.*

**Kata kunci:** *prediksi, kelompok, segmentasi pasar, K-means.*

### I. PENDAHULUAN

Segmentasi pasar adalah pengelompokan pasar untuk produk tertentu memprediksi target pasar, dimana setiap kelompok dapat dipilih sebagai pasar yang dituju sebagai target pemasaran suatu produk. [1]. Segmentasi pasar saat ini dijadikan beberapa perusahaan sebagai alat untuk mengukur hasil kinerja perusahaan terhadap tingkat kepuasan pelanggan yang bertransaksi. Setiap perusahaan diwajibkan mampu membagi atau mengelompokan pelanggan sesuai dengan karakteristik masing-masing dalam hal ini *potential customer* [2]. Analisis dan pengelompokan segmentasi pasar yang baik akan membawa perusahaan mendapatkan keuntungan kompetitif (*competitive advantages*) terhadap saingannya melalui diferensiasi produk. Perusahaan dapat membuat desain

produk dengan konsep yang memang menjadi kebutuhan dari konsumen itu sendiri, baik dari faktor kemasan, promosi produk, maupun konsep produk lainnya. Diferensiasi produk yang dibuat tidak akan mendapatkan keuntungan yang maksimal tanpa didukung dengan analisis faktor penyebaran produk tersebut, yaitu waktu penyebaran, identifikasi pelanggan yang potensial dan jumlah penjualan yang telah dilakukan. Semua faktor tersebut sebenarnya telah ada di banyak perusahaan dimana semuanya tercatat atau tersimpan di laporan penjualan maupun laporan distribusi produk yang telah dilakukan, namun banyak perusahaan menggunakan laporan atau data tersebut hanya sebatas sumber evaluasi tanpa adanya pengelolaan data yang lebih lanjut seperti pemanfaatan proses *data mining*.

Data yang ada seharusnya dapat dianalisis dan dimanfaatkan lebih lanjut dengan menggunakan metode *clustering* seperti penerapan metode K-means di setiap pengelompokan data yang ada. Analisis ini memberikan banyak *output* yang dapat dimanfaatkan perusahaan seperti pengelompokan penjualan produk dalam variabel waktu penjualan ataupun pengelompokan lainnya. Semua *output* tersebut dapat digunakan perusahaan untuk memenangkan persaingan ataupun meningkatkan *income* dan *omzet* perusahaan itu tersebut.

K-means merupakan bagian dari *metode* data *non-hierarchical clustering* dimana *metode ini* membutuhkan parameter-parameter batasan berdasarkan karakteristik pada masing-masing kelompok untuk mengklasifikasikan tingkat kemiripan dengan penentuan titik pusat awal sangat berpengaruh terhadap jumlah kelompok yang terbentuk dan

*Prosiding*  
**ANNUAL RESEARCH SEMINAR 2016**  
6 Desember 2016, Vol 2 No. 1

ISBN : 979-587-626-0 | UNSRI

<http://ars.ilkom.unsri.ac.id>

pengelompokan diukur dengan kedekatan objek sehingga setiap kelompok berisi data yang semirip mungkin terhadap nilai *mean* pada *cluster* atau pusat massa [3]. Perusahaan dapat memanfaatkan data yang ada dengan menggunakan algoritma K-means untuk membuat suatu pengelompokan dengan berfokus pada segmentasi, seperti segmentasi pasar, segmentasi *customer profiling* dan *segmentasi database*.

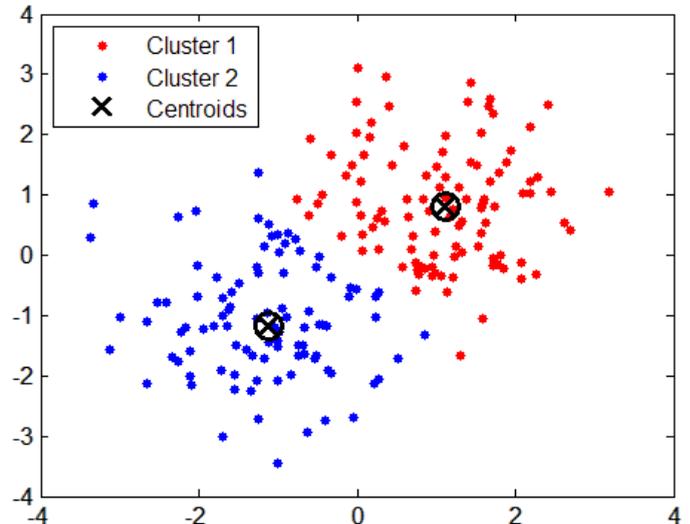
Hasil prediksi terbentuk sebanyak 3 *cluster* segmentasi pasar, yaitu *cluster* tinggi 36%, *cluster* sedang dan *cluster* rendah dengan persentase yang sama sebesar 32%.

## II. METODE PENELITIAN

Pada penelitian ini algoritma K-means meminimumkan total dari jarak anggota antar kelompok (jarak antara suatu anggota pada suatu kelompok dengan nilai *centroid cluster*). Similaritas anggota terhadap kelompok diukur dengan kedekatan objek terhadap nilai rata-rata pada kelompok. Untuk melakukan pengelompokan ini, nilai *k* harus ditentukan terlebih dahulu kelompok tersebut dengan suatu nilai tengah atau nilai pusat.

Langkah pengelompokan dengan algoritma K-means adalah sebagai berikut:

1. Memilih nilai *k* dengan Jumlah kelompok (*cluster*) yang diinginkan untuk menjadi titik yang merepresentasikan objek pada setiap kelompok.
2. Menetapkan inisialisasi *k* pusat *cluster* secara *random*.
3. Jika semua objek sudah dikelompokkan tempatkan setiap data atau objek ke kelompok terdekat. Kedekatan dua objek ditentukan berdasar jarak.
4. Menetapkan kelompok dengan keanggotaan *item* yang sekarang. Pusat kelompok adalah rata-rata dari semua data atau *objek* dalam *cluster* sampai kriteria konvergensi terpenuhi. Ilustrasi Algoritma K-means dapat diperlihatkan pada Gambar 1.

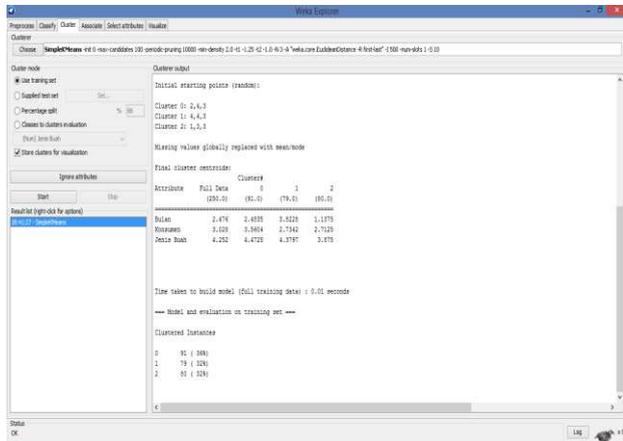


Gambar 1. Ilustrasi Algoritma K-means

Jumlah kelompok di atas merupakan percobaan yang baik karena semakin besar nilai maka semakin baik hasil *clustering* (pengklasteran) yang dilakukan. Dengan menambah jumlah *cluster* lebih dari 5 serta penggunaan *seed* yang berbeda-beda agar hasil yang didapat menjadi lebih akurat.

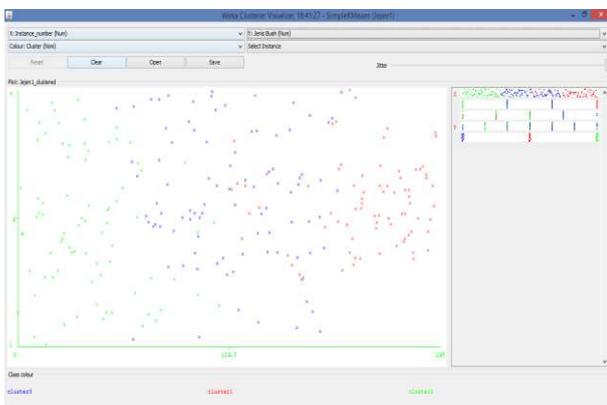
## III. ANALISIS DAN PEMBAHASAN

Untuk membuktikan apakah algoritma K-means dapat digunakan sebagai salah satu alat memprediksi kebutuhan buah maka diambil beberapa *sample* data sebagai pembuktian antara lain data penjualan buah dengan 3 *cluster* antara lain *cluster* tinggi, sedang dan rendah. Masing-masing *cluster* terdapat beberapa kategori sebagai alat perimbangan hasil analisis antara pangsa pasar penjualan buah, waktu penjualan buah dan jenis buah yang dijual. Adapun hasil pembahasan dengan menggunakan data tersebut dapat dilihat pada Gambar 2 dan Gambar 3.



Gambar 2. Hasil Persentasi Cluster

Gambar 2 menunjukkan penentuan jumlah cluster dan persentase cluster dari setiap objek ke cluster terdekat berdasarkan nilai centroid yang paling dekat selisihnya (jaraknya).



Gambar 3. Penyebaran hasil K-means

Gambar 3 menunjukkan visualisasi scatter plot pada nilai objek yang terdapat pada masing masing kelompok dan seed yang telah ditentukan. Berdasarkan hasil perhitungan K-means pada Gambar 2 di atas, maka hasil 3 kelompok sebagai berikut.

1. kelompok Tinggi dengan nilai persentase 36% dimiliki oleh jenis buah apel pada saat bulan Febuari dan pangsa pasar kuto yang paling diminati.
2. kelompok sedang dengan nilai persentase 32% dimiliki oleh jenis buah jeruk pada saat bulan Maret dan pangsa pasar Lemabang yang paling diminati menjadi kelompok sedang.
3. kelompok rendah dengan 32% rincian pada pasar perum dengan penjualan buah anggur pada bulan April menjadi kelompok rendah.

Dilihat dari penjelasan di atas, maka dapat disimpulkan bahwa algoritma K-means bisa digunakan untuk memprediksi hasil penjualan buah dengan hasil kelompok dan masing-masing persentase seperti Gambar 2.

#### IV. KESIMPULAN

Hasil pengujian sementara yang telah dilakukan penentuan centroid (titik pusat) pada tahap awal Algoritma K-means sangat berpengaruh pada hasil kelompok seperti pada hasil pengujian untuk mengklasifikasikan tingkat kemiripan dengan anggota terhadap kelompok diukur dengan kedekatan objek sehingga dalam setiap kelompok berisi data yang semirip mungkin terhadap nilai mean pada kelompok yang dilakukan dengan menggunakan dataset untuk pengelompokan data dalam jumlah yang banyak namun belum menghasilkan pengelompokan yang tepat. Penelitian selanjutnya akan memperbanyak jumlah data dan mengkaji secara mendalam faktor penyebab akurasi yang terbentuk.

#### REFERENSI

- [1]. *Determining the number of clusters in a data set.* [http://en.wikipedia.org/wiki/determining\\_the\\_number\\_of\\_clusters\\_in\\_a\\_data\\_set](http://en.wikipedia.org/wiki/determining_the_number_of_clusters_in_a_data_set).
- [2] Kasali, Rhenal, 2003, *Budaya Perusahaan*, Erlangga, Jakarta.
- [3] Ken, 2009, *Clustering Analysis, Part II: K-Means Clustering.* <http://www.centerspace.net>
- [4]. Mdarmayanti, Diah dan Tania Debby, 2014, *Market Segmentation, Targeting, dan Brand Positioning* dari *winston Premier*, Universitas Kristen Petra, Surabaya.
- [5]. Nango Dwi Novianti, 2012, Penerapan Algoritma K-means untuk clustering data anggaran pendapatan belanja daerah di kabupaten XYZ, Universitas Negeri Gorontalo.