

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi internet saat ini begitu pesat dalam berbagai bidang, tak terkecuali dalam dunia bisnis. hal ini dapat dilihat dengan munculnya berbagai usaha dibidang penjualan berbasis online atau biasa disebut *electronic commerce*. *E-Commerce* memberikan manfaat bagi perusahaan sehingga akan menjadi keunggulan kompetitif bagi perusahaan yang mengaplikasikannya. (Rudy, Wahyudiarti, Megaputri, & Wihardini, 2008). Lembaga atau perusahaan yang mengaplikasikan *E-Commerce* dalam kegiatan pemasaran atau perdagangannya biasa dikenal dengan nama *Online shop*. Kegiatan pemasaran pada sistem ini lebih praktis dan efisien tanpa adanya interaksi langsung antara pelaku usaha dan konsumen dan akan memperlihatkan kemudahan bertransaksi dibandingkan proses pemasaran pada toko konvensional.

Semakin besar sebuah *Online shop* maka semakin banyak transaksi yang dilakukan, serta dapat menarik data yang begitu besar pula. Data tersebut dapat berupa data konsumen, data penjualan dan masih banyak lagi. Tapi terkadang dari besarnya data yang diperoleh, Perusahaan tidak sadar betapa berharganya data tersebut dan hanya dipandang sebagai arsip yang sudah menjadi berkas yang tidak terpakai. Hal tersebut merupakan pandangan yang salah dikarenakan jika mampu mengolahnya dengan tepat, akan dapat memberikan berbagai informasi yang sangat bermanfaat bagi perusahaan.

Salah satu informasi yang dapat diperoleh yaitu untuk mengetahui minat konsumen pada penjualan produk, dimana minat konsumen terhadap penjualan suatu produk dapat diukur dari banyaknya jumlah transaksi penjualan yang dilakukan. Dalam sebuah transaksi penjualan minat konsumen dapat dibedakan menjadi beberapa kelompok berdasarkan tingkatannya. Maka teknik pengolahan data sangat diperlukan untuk menggali informasi dari data tersebut. Dalam hal ini maka digunakanlah teknik *Data Mining*.

Salah satu metode dalam *Data Mining* yaitu *Clustering* atau pengelompokan. Dari beberapa teknik *Clustering* yang paling sederhana dan umum dikenal adalah algoritma *K-Means*. *K-Means* merupakan salah satu metode data *clustering non hirarki* yang mempartisi data ke dalam *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok lain (Agusta, 2007).

Berdasarkan permasalahan tersebut maka penulis mengambil judul “Implementasi Algoritma K-Means *Clustering* untuk Pengelompokan Minat Konsumen pada Produk Online Shop”.

## 1.2 Rumusan Masalah

Berdasarkan pemaparan latar belakang diatas maka permasalahan yang dapat diambil adalah :

1. Berapa jumlah kelompok yang paling optimal untuk data penjualan produk *Online Shop* ?
2. Berapa jumlah data pada masing-masing *Cluster* yang terbentuk ?

## 1.3 Batasan Masalah

Ruang lingkup batasan masalah yang dibahas adalah :

1. Dataset yang digunakan adalah dataset penjualan produk online shop yang diperoleh dari *UCI Machine Learning Repository*.
2. File yang digunakan adalah file *excel* berformat *.xls* atau *.xlsx*.
3. Input sistem menggunakan data penjualan produk dalam 1 bulan periode transaksi.
4. Atribut yang digunakan yaitu jumlah transaksi, total penjualan dan rata-rata penjualan.
5. Pengukuran performa menggunakan aplikasi *rapidminer studio*.
6. Penerapan sistem menggunakan aplikasi berbasis *web (php)*.
7. Database sistem menggunakan *MySQL*.

#### 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Mencari kelompok atau jumlah *Cluster* paling optimal untuk data penjualan produk *Online Shop*.
2. Mencari jumlah data dalam *Cluster* pada masing-masing *Cluster* yang terbentuk.

#### 1.5 Manfaat Penelitian

Adapun manfaat penelitian ini adalah :

1. Membantu perusahaan dalam mengetahui informasi tingkat minat konsumen terhadap produk yang dipasarkan.
2. Mempercepat proses pengambilan keputusan dalam penetapan strategi pemasaran suatu produk.

