

# **SISTEM PENGOALAHAN DATA MINING INDUSTRI SEPATU MENGGUNAKAN METODE K-MEANS CLUSTERING DI JAWA TENGAH**

Ahmad Zaqi Al Kahfi  
Universitas Dian Nuswantoro  
Email : fawkeszach@gmail.com

## **Abstrak**

Pengolahan database dalam dunia industri di Jawa Tengah masih mengalami kesulitan. Data-data tersebut dalam jumlah yang banyak dan memiliki nilai ukuran yang tidak tetap. Data mining adalah teknik yang digunakan untuk mengekstrak pola dari data mentah yang sebenarnya tidak terlalu berguna dan memiliki ukuran yang besar agar dapat dianalisa. Dalam data mining terdapat banyak metode pengolahan data yang dapat digunakan, salah satunya adalah dengan menggunakan metode *K-Means Clustering* yaitu sebuah algoritma untuk mengklasifikasikan atau mengelompokkan objek-objek. Data Industri Kecil Menengah (IKM ) Sepatu di wilayah Jawa Tengah. Pengolahan data pada Dinas Perindustrian dan Perdagangan (DINPERINDAG) Povinsi Jawa Tengah khususnya dari IKM Sepatu masih dalam bentuk yang belum pengolahan dengan baik.

Kata kunci : Data Mining, K-means, IKM, Dinperindag

### **1. PENDAHULUAN**

Dinperindag untuk mengambil kebijakan.

#### **1.1 Latar Belakang Masalah**

Saat ini pengolahan database dalam dunia industri di Jawa Tengah masih mengalami kesulitan.

#### **1.3 Batasan Masalah**

Pembuatan sistem pengolahan data hanya sebatas pada IKM Sepatu di wilayah provinsi Jawa Tengah. Sistem yang dibuat adalah system yang digunakan untuk mengklasifikasikan data IKM Sepatu di Jawa Tengah. Sistem yang dibuat tidak meliputi keamanan untuk melindunginya dari serangan pihak yang tidak bertanggung jawab

#### **1.2 Rumusan Masalah**

Data belum dapat diklasifikasikan menjadi bagian-bagian sesuai dengan kebutuhan. Pengelompokkan dan pengklasifikasian data masih dilakukan secara manual dan tidak objektif. Informasi dari data IKM Sepatu belum dapat membantu

#### **1.4 Tujuan Penelitian**

Membuat aplikasi dengan teknik pengolahan *data mining* untuk mengklasifikasikan data IKM Sepatu se-Provinsi Jawa Tengah. Membuat aplikasi agar dalam mengelompokkan data tidak dilakukan secara manual dan objektif. Membuat aplikasi agar memudahkan pemerintah dalam memperoleh informasi demi membuat kebijakan. Membuat aplikasi dengan interface yang baik agar informasi yang disampaikan mudah dimengerti dan dianalisis.

#### **1.5 Manfaat Penelitian**

Membantu Dinas Perindustrian dan Perdagangan dalam mengolah data IKM Sepatu se-Provinsi Jawa Tengah. Membantu pemerintah dalam memperoleh informasi demi terciptanya kebijakan yang sesuai dengan kebutuhan pelaku industry dan pedagang di Jawa Tengah. Membantu meningkatkan produktifitas IKM sepatu di provinsi Jawa Tengah. Membantu pemerintah dalam memahami dan menganalisis informasi yang ada dengan cara memvisualisasikannya dalam bentuk grafik.

## **2. Landasan Teori**

### **2.1 Data Mining**

Data mining adalah ekstraksi pola yang menarik dari data dalam jumlah besar. Dengan kata lain, data mining merupakan rangkaian proses untuk menghasilkan suatu informasi yang bernilai dari sekumpulan data yang tidak bernilai. Jadi data mining juga bisa dikatakan sebagai pengolahan data menjadi suatu informasi yang memberikan pengetahuan

### **2.2 Metode Data Mining**

- Association Rule Learning  
Analisis asosiasi atau association rule mining adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item
- Clustering  
Clustering adalah metode penganalisaan data, yang sering dimasukkan sebagai salah satu metode Data Mining, yang tujuannya adalah untuk mengelompokkan data dengan karakteristik yang sama ke suatu 'wilayah' yang sama dan data dengan karakteristik yang berbeda ke 'wilayah' yang lain.
- Decision tree  
Dalam decision tree tidak menggunakan vector jarak untuk mengklasifikasikan obyek. Seringkali data observasi mempunyai atribut-atribut yang bernilai nominal.

### 2.3 K-means Clustering

*K-Means* termasuk dalam *partitioning clustering* yaitu setiap data harus masuk dalam *cluster* tertentu dan memungkinkan bagi setiap data yang termasuk dalam *cluster* tertentu pada suatu tahapan proses, pada tahapan berikutnya berpindah ke *cluster* yang lain. K-Means memisahkan data ke  $k$  daerah bagian yang terpisah, dimana  $k$  adalah bilangan integer positif.

$$d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - c_{kj})^2} \quad (\text{Rumus 2.1})$$

2.1)

$$\text{Min } \sum_{k=1}^k d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - c_{kj})^2}$$

(Rumus 2.2)

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad (\text{Rumus 2.3})$$

Algoritma dasar dalam K-Means adalah :

1. Tentukan jumlah kluster ( $k$ ), tetapkan pusat kluster sembarang.

2. Hitung jarak setiap data ke pusat kluster menggunakan rumus 3.1.
3. Kelompokkan data ke dalam kluster yang dengan jarak yang paling pendek menggunakan rumus 3.2.
4. Hitung pusat kluster yang baru menggunakan rumus 3.3.

Ulangi langkah 2 sampai dengan 4 hingga sudah tidak ada lagi data yang berpindah ke kluster yang lain.

### 3. Metodologi Penelitian

#### 3.1. Obyek Penelitian

Dalam menyusun Tugas Akhir penulis mengambil obyek penelitian pada Industri Kecil Menengah (IKM) Sepatu di Jawa Tengah dengan menggunakan bahasa pemrograman php & mySQL.

### **3.2. Tahap-tahap Pengembangan Sistem**

Dalam perancangan dan pengembangan tugas akhir ini, penulis menggunakan metode pengembangan sistem *waterfall*. Pembuatan aplikasi pembelajaran tentang sistem peredaran darah dengan metode pengembangan ini, disusun dengan lima tahapan. Lima tahapan itu adalah analisa kebutuhan sistem (*requirements*), tahap perancangan (*design*), tahap implementasi dari perancangan yang sudah dibuat (*implementation*), tahap uji coba atau pengujian (*verification*) dan tahap pemeliharaan (*maintenance*).

#### **3.2.1 Perancangan Alur Sistem**

Tahap perancangan sistem dimulai setelah tahap analisa selesai dan didefinisikan secara jelas. Di dalam tahap ini akan dijelaskan lebih rinci lagi dengan didasarkan pada tahapan

sebelumnya, kemudian detail aliran proses dari software yang akan dibangun. Hasil dari tahap perancangan dapat diterapkan menjadi prosedur-prosedur dengan alat bantu bahasa pemrograman yang ada.

#### **3.2.2 Diagram Alir Data/Data Flow Diagram (DFD)**

Diagram alir data merupakan penggambaran lebih detail dan lebih rinci dari proses yang telah digambarkan sebelumnya pada konteks diagram. Dari proses yang dijelaskan pada konteks diagram akan diperinci ke DFD level 0, dan proses yang dijelaskan pada DFD level 0, akan diperinci lagi pada DFD level 1.

### **4. Hasil dan Pembahasan**

#### **4.1 Analisa Sistem Saat Ini**

Saat ini system sudah di operasikan dan dimanfaatkan oleh DINPERINDAG sebagai media pembantu kinerja pengolahan data industry sepatu yang ada di Jawa Tengah. Dalam pengoperasian

system berjalan dengan lancar. Setiap kebutuhan sistem yang tersedia berjalan dengan baik. Keamanan system dalam menjaga data – data yang ada di proteksi dengan ketat. Terdapat beberapa jalur akses system yang dikategorikan menjadi tiga macam, yaitu akses sebagai admin super, admin biasa atau sebagai user.

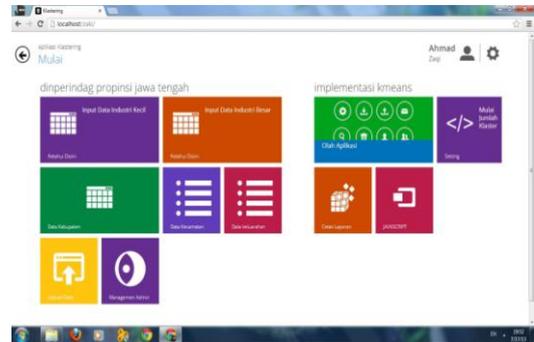
## 4.2 Tampilan Program

### Halaman Login



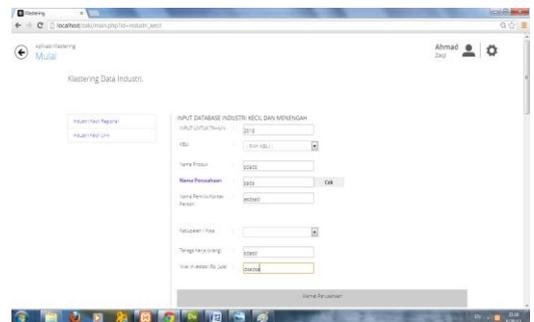
Gambar di atas merupakan tampilan login dari Sistem informasi statistik industri Se-Jawa Tengah di Dinas Perindustrian dan Perdagangan Provinsi Jawa.

### Halaman Utama



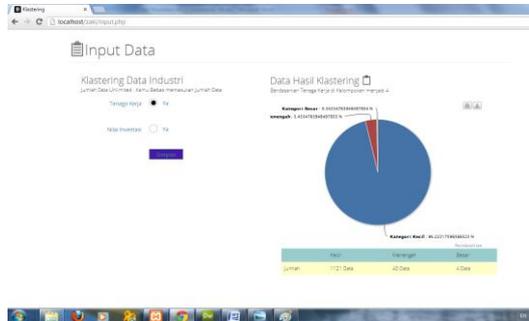
Gambar diatas merupakan Halaman Utama pengolahan *Clustering*, terdiri dari beberapa bagian pada kolom sebelah kiri berfungsi untuk menginput data dan mencetak laporan sedangkan kolom sebelah kanan untuk pengolahan *clustering* data.

### Halaman Untuk Input Data



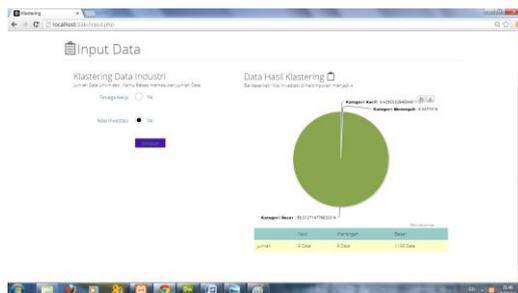
Pada halaman ini ,Admin Dinas Provinsi mempunyai hak akses untuk mengelola nama inputan data Kabupaten, Kecamatan, Kelurahan dan KBLI yang semua saling terkait seperti yang sudah di jelaskan di gambar relasi database.

## Halaman Clustering Data Industri Berdasarkan Tenaga Kerja.



Pengklasifikasian data industri berdasarkan tenaga kerja kemudian memberikan aksi klik pada sistem dan sistem merespon data industri sesuai dengan setting kluster dan data industri yang ada sehingga menghasilkan graph tertentu.

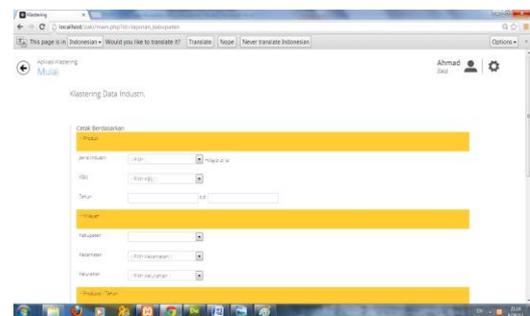
## Halaman Clustering Data Industri Berdasarkan Nilai Investasi.



Pengklasifikasian data industri berdasarkan nilai investasi kemudian memberikan aksi klik pada sistem dan sistem merespon data industri sesuai dengan setting kluster dan

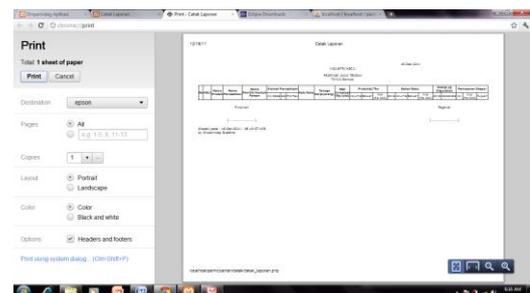
data industri yang ada sehingga menghasilkan graph tertentu.

## Halaman Cetak Laporan



Gambar diatas adalah tampilan halaman untuk Mencetak Laporan dari Kabupaten tertentu. Halaman tersebut hanya dapat di akses oleh Admin Dinas Provinsi yang berhak mencetak Laporan dari tiap kabupaten nya.

## Halaman Cetakan Laporan



User dapat Langsung Mencetak Laporan menggunakan Printer seperti yang ada atau jika ingin mendownload nya User dapat memilih tombol Download.

## 5. Kesimpulan

Dari hasil aplikasi Sistem Informasi Statistik Industri dan pemetaan nya di Dinas Perindustrian dan Perdagangan Provinsi Jawa tengah dapat penulis simpulkan sebagai berikut :

- a. Aplikasi yang di buat oleh penulis ini dapat membantu Sub.bagian Program untuk mengolah data menjadi laporan.
- b. Aplikasi ini membantu Masyarakat yang membutuhkan data Industri untuk mengambil Informasi secara lengkap dengan data yang akurat.
- c. Adanya database yang terintegrasi pada aplikasi ini sehingga informasi yang di berikan lebih valid .
- d. Aplikasi ini dapat membantu bagian Pengolahan Industri dalam mengolah Industri yang ada dengan meminimalisir redudansi data.

## Daftar Pustaka

- [1] Kardi.2007.*K-meansClusteringTutorial*, <http://people.revoledu.com/kardi/index.html>).
- [2] Agusta, Yudi.Pebruari 2007 . “ *KMeans Penerapan, Permasalahan dan Metode Terkait*”. Jurnal Sistem dan Informatika Vol.3 : 47-60.
- [3] Han,J.and Kamber,M,2006, “ *Data Mining Concepts and Techniques Second Edition*” . Morgan Kauffman, San Francisco.
- [4] Davies, and Paul Beynon, 2004, “*Database Systems Third Edition*” , Palgrave Macmillan, New York.
- [5] Pramudiono, I. 2007. *Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data*. <http://www.ilmukomputer.org/wpcontent/uploads/2006/08/iko-datamining.zip>

- Diakses pada tanggal 10 April 2013 jam 09.46.
- [6] Pramudiono, I., 2007, *Algoritma Apriori*, <http://datamining.japati.net/cgi-bin/indodm.cgi?bacaarsip&1172210143> Diakses pada tanggal 11 April 2013 jam 11.03
- [7] Santosa, Budi, 2007, “*Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*”, Graha Ilmu, Yogyakarta.
- [8] Bramer, Max , 2007, “*Principles of Data Mining*”, Springer, London.
- [9] Elmasri, Ramez and Shamkant B. Navathe, 2000, “*Fundamentals of Database Systems. Third Edition*”, Addison Wesley Publishing Company, New York.
- [10] Kadir, Abdul, 1999, “*Konsep dan Teknik Praktis Basis Data*”, Penerbit Andi, Yogyakarta.
- [11] Pressman, Roger S, 1997, “*Software Engineering: A Practitioner's Approach*.” The McGraw-Hill Companies, Inc., New York.
- [12] Sommerville, Ian, 2003, “*Software Engineering (Rekayasa Perangkat Lunak)/ Edisi 6/Jilid 1*” Erlangga, Jakarta.
- [13] Witten, I. H and Frank, E. 2005. *Data Mining : Practical Machine Learning Tools and Techniques Second Edition*. Morgan Kauffman : San Francisco.
- [14] Chintakayala, Padmini. 2005. “*Beginners Guide for Software Testing : Symbiosys Technologies*”.