

**ETS E-Bisnis (E)**

**ANALISA K-MEANS CLUSTERS**



**Mochammad Iqbal Rizky**

**1211800229**

**E-BISNIS**

**Fakultas Ekonomi dan Bisnis**

**Program Studi Manajemen**

**Universitas 17 Agustus 1945 Surabaya**

**2020**

## **Kata Pengantar**

Kemajuan teknologi membuat kita lebih mudah mencari informasi dimanapun kapanpun karena semakin tahun selalu ada inovasi teknologi setiap tahun. Oleh karena itu pentingnya kita belajar tentang teknologi supaya kita bisa beradaptasi dan tidak tertinggal jaman

Dalam makalah ini, diuraikan mengenai bagaimana teknolog di era modern ini. Semoga dengan uraian mengenai Analisis K-Means Clusters sekarang ini, maka pembaca bisa mendapat gambaran yang lebih jelas mengenai tema terkait. Penulis terbuka bila para pembaca ada yang memiliki saran atau masukan terhadap makalah yang ditulis.

## **DAFTAR ISI**

**COVER**

**KATA PENGANTAR**

**DAFTAR ISI**

**BAB I PENDAHULUAN**

**1. RUMUSAN MASALAH**

**BAB II TINJAUAN PUSTAKA**

**1. DATA MINING**

**BAB III ANALISIS WEB SRIBULANCER**

**1. K-MEANS**

**2. PENGENALAN POLA**

**3. TAHAP TAHAP DATA MINING**

**BAB IV PENUTUP**

**1. KESIMPULAN**

**DAFTAR PUSTAKA**

## **BAB 1 PENDAHULUAN**

### **Latar Belakang**

Perkembangan Teknologi mulai pesat. Kemajuan teknologi membuat kita lebih mudah mencari informasi dimanapun kapanpun karena semakin tahun selalu ada inovasi teknologi setiap tahun. Oleh karena itu pentingnya kita belajar tentang teknologi supaya kita bisa beradaptasi dan tidak tertinggal jaman

Makadari itu inovasi K-Means Clusters sangat membantu dalam mencari data untuk memudahkan dan lebih efisien

### **Rumusan Masalah**

Berdasarkan faktor-faktor yang telah kami uraikan diatas, kami akan melakukan analisis terhadap K-Means

## **BAB II Tinjauan Pustaka**

### **Data Mining**

#### **1. Data Mining**

Data Mining adalah sebuah proses mencari pola, hubungan, penggalian nilai tambah dari data dan informasi yang berukuran besar berupa pengetahuan dengan tujuan menemukan hubungan dan menyederhanakan data agar diperoleh informasi yang mudah dipahami dan bermanfaat

Menurut Davies Data mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar

Menurut Pramudiono Data mining juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data

Berdasarkan beberapa definisi karya di atas, peneliti dapat menarik kesimpulan tentang makna karya. Data mining adalah suatu kegiatan yang dilakukan seseorang untuk menggali atau menambang informasi, untuk tujuan menemukan hubungan dan menyederhanakan data supaya mudah untuk dipahami

## **BAB III Analisis K Means**

### **1. K Means Clusters**

K-Means cluster adalah algoritma data mining atau penggalian data dengan kemampuan mengelompokkan data. K-Means untuk menghasilkan informasi baru berupa klasifikasi kelompok produk berdasarkan data transaksi penjualan di masa lampau

### **2. Pengenalan Pola**

Menurut Santoso Pengenalan Pola, Data Mining, dan Machine Learning Pengenalan pola adalah suatu disiplin ilmu yang mempelajari cara-cara mengklasifikasikan obyek ke beberapa kelas atau kategori dan mengenali kecenderungan data. Tergantung pada aplikasinya, obyek-obyek ini bisa berupa pasien, mahasiswa, pemohon kredit, image atau signal atau pengukuran lain yang perlu diklasifikasikan atau dicari fungsi regresinya. Data mining, sering juga disebut knowledge discovery in database (KDD), adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan. Sehingga istilah pattern recognition jarang digunakan karena termasuk bagian dari data mining

### **3. Tahap-tahap data mining**

1. Pembersihan data (data cleaning) Pembersihan data merupakan proses menghilangkan noisedan data yang tidak konsisten atau data tidak relevan. Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa data mining yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi performansi dari teknik data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.
2. Integrasi data (data integration) Integrasidata merupakan penggabungan data dari berbagai database ke dalam satu database baru. Tidak jarang data yang diperlukan untuk data mining tidak hanya berasal dari satu database tetapi juga berasal dari beberapa database atau file teks. Integrasi data dilakukan pada atribut-aribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang

menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

3. Seleksi Data (Data Selection) Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus market basket analysis, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja
4. Transformasi data (Data Transformation) Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan clustering hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi- bagi menjadi beberapa interval. Proses ini sering disebut transformasi data.
5. Proses mining, Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.
6. Evaluasi pola (pattern evaluation), Untuk mengidentifikasi pola-pola menarik kedalam knowledge based yang ditemukan. Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba metode data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.
7. Presentasi pengetahuan (knowledge presentation), Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses data mining. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data

## **BAB IV PENUTUP**

### **KESIMPULAN**

Dari semua data analisis ini bahwa di era modern ini semakin mudah mencari data banyak sekali dan tidak terbatas. Dengan adanya K-Means Clusters mencari informasi lebih efisien. Akan tetapi semua teknologi juga ada kelebihan dan kekurangan masing-masing

## DAFTAR PUSTAKA

- [1] S. Supangat, "Pemanfaatan Fitur Analisis Data Menggunakan K-Means Cluster Dalam Point of Sales (POS) *P3M*, Institut Informatika Indonesia (IKADO) Surabaya,. 2019, Accessed: 2019/11/31. [Online]. Available <http://ejournal.ikado.ac.id/index.php/teknika/article/view/157>