

# **IMPLEMENTASI ALGORITMA *K-MEANS*++ UNTUK *CLUSTERING* PENJUALAN BAHAN BANGUNAN (“STUDI KASUS UD SUMBER BANGUNAN”)**

## **BAB I PENDAHULUAN**

### **1.1 Latar Belakang**

Di era industri 4.0 kemajuan dibidang teknologi informasi dan komunikasi saat ini berkembang pesat disegala bidang seperti pada bidang industri, ekonomi, pendidikan, konstruksi dan lain sebagainya. Sehingga dengan kemajuan teknologi manusia dimudahkan dalam hal mencari informasi yang ingin didapatkan. Teknologi informasi sangat dibutuhkan untuk pelaku usaha sebagai tolak ukur dalam memutuskan keputusan yang dapat memberikan dampak bagi keberlangsungan bisnis. Kita dituntut untuk senantiasa mengembangkan bisnis agar selalu bertahan dalam persaingan, dalam pengembangan bisnis dibutuhkan sebuah strategi untuk meningkatkan kualitas penjualan seperti meningkatkan kualitas produk, penambahan jenis produk, dan pengurangan biaya operasional perusahaan dengan cara menggunakan analisis data perusahaan. Namun hal ini belum sepenuhnya dimanfaatkan oleh pemilik UD Sumber Bangunan.

UD Sumber Bangunan merupakan sebuah usaha yang bergerak dalam bidang penjualan bahan bangunan, tetapi dalam hal bisnis UD Sumber Bangunan belum menggunakan komputer untuk menjalankan sistemnya, data-data yang terlibat dalam setiap transaksi pada toko ini, seperti data item yang dibeli untuk memenuhi stok, jumlah item yang dibeli oleh konsumen, data item yang terjual dan data konsumen. Karena banyaknya data yang hanya dibiarkan begitu saja dan hanya berguna sebagai arsip toko sehingga kerap timbul masalah seperti penumpukan stok dan kepuasan konsumen. Padahal dengan data yang ada, pemilik toko bisa memanfaatkan untuk strategi

pemasaran. Data-data tersebut bisa dijadikan sebagai alat dalam mengambil keputusan untuk solusi bisnis dengan dukungan sebuah teknologi data mining. Data mining adalah proses penemuan pola yang tidak diketahui sebelumnya dalam data besar dan kompleks (Jiawei Han, Micheline Kamber, dan Jian Pei 2011). Informasi berupa data transaksi yang besar membutuhkan sebuah sistem yang secara otomatis mencari hubungan antar item dalam data di Database. Dapat dicontohkan dari toko UD Sumber Bangunan dari data yang diperoleh dari pemilik toko.

Dalam penerapan algoritme *K-Means++* dan jarak pengukuran *Euclidean* dari kesamaan dipilih untuk digunakan dalam analisis nilai. Pada penelitian ini dalam menganalisa data mining digunakan algoritme *K-Means++ clustering*. *Clustering* merupakan salah satu teknik dari salah satu fungsionalitas data mining, Algoritme *Clustering* merupakan Algoritme pengelompokkan sejumlah data menjadi kelompok-kelompok data tertentu (*cluster*) (Parsaoran Tamba & Toknady Kesuma, 2019). Algoritme *K-Means* terkenal karena Algoritme *K-Means* relatif sederhana dan efisien untuk digunakan. *K-Means* berkerja dengan cara melakukan pengelompokkan secara partisi yang memisahkan data dalam kelompok-kelompok tertentu dengan meminimalkan rata-rata jarak setiap data klasternya (Triyandana et al., 2022).

*K-Means* merupakan salah satu metode untuk menganalisis data. Algoritme ini menentukan jumlah serta nilai *cluster* (*k*) secara acak. Nilai tersebut menjadi pusat awal dari *cluster* atau bisa disebut dengan *centroid* (Nasir, 2021a). Pemilihan *centroid* yang dilakukan secara acak, menjadi kelemahan dari proses tersebut. Oleh karena itu, digunakan metode *K-Means++* yang merupakan improvisasi dari *K-Means* untuk mengatasi masalah tersebut dengan memilih pusat *cluster* awal secara cermat. Tujuan utama dari Algoritme ini adalah untuk mengambil titik-titik data tersebut sebagai pusat awal yang berjarak sejauh mungkin antara satu sama lain (Sri Fastaf & Yamasari, 2022a).

Dengan menggunakan Algoritme ini, data-data yang telah didapatkan dapat dikelompokkan kedalam beberapa *cluster* berdasarkan kemiripan data tersebut, sehingga data-data yang memiliki kemiripan karakteristik yang sama dikelompokkan

dalam satu cluster dan yang memiliki karakteristik berbeda di kelompokkan dalam *cluster* yang lain yang memiliki karakteristik yang sama. Dengan dilakukannya pengelompokan-pengelompokan data seperti ini diharapkan pemilik usaha dapat menentukan strategi yang tepat untuk melayani kebutuhan konsumen

## 1.2 Rumusan Masalah

Berdasarkan uraian permasalahan yang dipaparkan dibagian latar belakang, maka rumusan permasalahan yang didapatkan antara lain:

1. Bagaimana menerapkan Algoritme *Clustering* yaitu *K-Means++* dan *K-Means* untuk data penjualan bangunan di Toko UD Sumber Bangunan?
2. Bagaimana hasil perbandingan Algoritme *Clustering* menggunakan *K-Means++* dan *K-Means* ?
3. Bagaimana melakukan rancang bangun Sistem *Clustering* penjualan Bahan Bangunan dengan *K-Means++* menggunakan Metode *Waterfall*?

## 1.3 Batasan Masalah

Berdasarkan rumusan masalah diatas, Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Sumber data yang digunakan berasal dari toko UD Sumber Bangunan pada bulan Januari 2023 - Februari 2023.
2. Algoritme *Clustering* yang digunakan adalah algoritme *K-Means* dan algoritme *K-Means++* untuk inialisasi centroid dengan model *Euclidean Distance*.
3. Output yang dihasilkan berupa clusterisasi hasil penjualan barang di toko UD Sumber Bangunan.
4. Hasil penelitian berfokus pada hasil *cluster* beberapa kelompok dan kesimpulan yang diambil dari hasil analisis.
5. Data atribut terdiri jumlah Stok Awal, Total Penjualan, dan Stok Akhir
6. Implementasi sistem saat input nilai hanya bisa menggunakan format xls.

#### 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan Algoritme *K-Means++ Clustering* dengan mengelompokkan barang-barang keberapa cluster berdasarkan kemiripan data tersebut.
2. Menganalisis dan memahami bagaimana penggunaan *K-Means++* memengaruhi hasil klaster yang dihasilkan oleh Algoritme *K-Means++*
3. Membantu memastikan sistem *Clustering* penjualan bahan bangunan yang dibangun berkualitas tinggi dan sesuai dengan kebutuhan pengguna.

#### 1.5 Manfaat Penelitian

Manfaat dari penelitian ini diharapkan dapat digunakan sebagai berikut :

1. Manfaat bagi objek: Pemilik usaha dapat menentukan strategi yang tepat untuk melayani kebutuhan konsumen.
2. Manfaat bagi peneliti: Menambah pengetahuan mengenai penerapan Algoritme *K-Means Clustering*, *K-Means++* dan *Euclidean Distance*.
3. Manfaat bagi pembaca: Dapat menambah wawasan bagi pembaca dan dapat digunakan sebagai referensi untuk penelitian selanjutnya.

#### 1.6 Metodologi Penelitian

Tahapan yang akan dilakukan dalam penelitian ini adalah sebagai berikut:

1. Pengumpulan Data

Proses pengambilan data ini dilakukan dengan observasi dilokasi penelitian yakni UD Sumber Bangunan. Selain observasi, pengumpulan data juga dilakukan dengan wawancara pemilik usaha.

2. Studi literatur

Proses studi literatur dimulai dengan pencarian topik penelitian yang berasal dari permasalahan yang ditemui di lingkungan sekitar. Setelah didapatkan topik yang diambil. Referensi tersebut dapat dijadikan acuan untuk menentukan variabel ataupun metode penelitian.

### 3. Tahap Perancangan

Tahap ini dilakukan untuk membuat data mentah yang akan diolah menjadi data yang berkualitas. Hal ini dilakukan agar dapat memperoleh hasil yang lebih akurat dalam pemakaian algoritme *K-Means++ Clustering*.

### 4. Analisis Sistem

Tahap ini dilakukan berdasarkan hasil observasi dan pengumpulan data yang dilakukan. Analisa kebutuhan sistem dilakukan untuk menentukan fitur yang digunakan dalam sistem.

### 5. Implementasi

Merupakan proses penerjemahan dari tahap perancangan ke dalam bentuk aplikasi dengan bahasa pemrograman *Javascript* dan *database* menggunakan MySQL.

### 6. Pengujian

Tahap ini dilakukan untuk melakukan uji coba terhadap program yang dibangun dan menguji sejauh mana kinerja sistem dan keakuratan metode sehingga dapat menghasilkan informasi yang diharapkan.

### 7. Tahap Penyusunan Laporan

Tahap ini dilakukan penyusunan laporan dari penelitian berdasarkan sistematika penulisan.

## 1.7 Sistematika Penulisan

Sistematika penulisan merupakan sebagai upaya untuk memudahkan pembaca dalam memahami materi skripsi secara integral dan dapat memenuhi persyaratan sebagai karya ilmiah, perlu dibuat suatu rangkaian relasi materi antar bab diorganisasikan sebagai suatu satuan yang utuh. Sehubungan dengan pemikiran itu, maka aliran materi per bab dalam skripsi disusun secara sistematis dan runtut melalui sistematika penulisan sebagai berikut:

## **BAB I PENDAHULUAN**

Bab ini merupakan pendahuluan dari penulisan yang berisikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

## **BAB II LANDASAN TEORI**

Landasan teori berisi uraian sistematis dari teori yang ada pada literatur maupun penjabaran tinjauan pustaka yang mendasari pemecahan masalah serta langkah-langkah pengolahan data yang diperoleh dari hasil penelitian dan juga konsep-konsep yang dijadikan kerangka dan landasan berfikir dalam pengembangan model penelitian ini, serta langkah-langkah yang ditempuh dalam proses penelitian masalah.

## **BAB III ANALISIS DAN PERANCANGAN SISTEM**

Pada bab ini akan dijelaskan tentang definisi sistem, perangkat keras dan perangkat lunak yang digunakan, relasi antar tabel, diagram alir sistem, rancangan masukan dan keluaran.

## **BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM**

Pada bab ini akan dijelaskan tentang manfaat sistem dan implementasinya, sehingga dapat menghasilkan informasi sesuai dengan kebutuhan.

## **BAB V PENUTUP**

Pada bab ini berisi kesimpulan dan saran-saran pengembangan penelitian yang akan datang.

