

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada masa yang serba cepat ini, data adalah yang paling dicari dalam semua aspek kehidupan. Data-data yang dicari harus terkini, akurat, dan bisa dipercaya. Untuk itu dibutuhkan dukungan teknologi di dalamnya untuk mengelola data-data tersebut menjadi informasi dengan cepat, tepat, dan efisien. Pada masa yang serba otomatis ini semua orang berlomba untuk memperbarui teknologinya untuk meningkatkan pemrosesan data guna mempercepat akses informasi. Perkembangan pemrosesan data ini juga berdampak dalam bidang bisnis secara signifikan. Khususnya dengan bagaimana perubahan cara bisnis berjalan. Dampak yang paling nyata terjadi adalah bagaimana dahulu data diolah dan disimpan secara manual dan kemudian pada masa sekarang semuanya terkomputerisasi (Adriana, 2024).

Penggunaan data yang banyak yang kemudian diolah dengan komputer sangat dibutuhkan dalam dunia bisnis. Hal tersebut juga sejalan dengan penelitian dari (Damanik & Nasution, 2024) dengan judul “Evaluasi Penggunaan Teknologi Big Data untuk Analisis Data Bisnis dan Pengambilan Keputusan” yang menunjukkan pada dunia bisnis banyaknya jumlah data akan mengikuti dengan skala bisnis yang membutuhkan. Maka semakin besar skala bisnis suatu usaha maka jumlah data beredar dan data yang dibutuhkan akan semakin besar. Dengan besarnya jumlah data yang digunakan maka dibutuhkan pengelolaan data yang efisien dan cepat untuk mendapatkan informasi. Dalam dunia bisnis data adalah alat yang efektif untuk banyak hal, seperti mengidentifikasi pola, menganalisis tren pasar, maupun mengetahui keinginan konsumen secara mendetail. Seluruh analisis itu digunakan untuk membantu dalam pengambilan keputusan.

Banyak perusahaan yang membutuhkan pengelolaan data yang baik untuk membantu pengambilan keputusan di masa depan. Salah satu bentuk pengelolaannya dengan mengelompokkan dengan baik dan benar. Pengelompokan data atau *Data Clustering* merupakan teknik dalam data *mining* yang bertujuan untuk mengelompokkan data berdasarkan kesamaan pada karakteristik tertentu.

Data yang sudah dikelompokan akan dianalisis untuk mengungkap pola-pola yang tersembunyi atau hubungan yang signifikan antar data (Hendrastuty, 2024).

Terdapat berbagai jenis metode dalam mengelompokan data, antara lain *Hierarchical Clustering* yaitu metode pengelompokan data dengan membangun pohon *cluster* yang disebut dengan *dendrogram*. Setiap *node* dalam *dendrogram* berisi klaster anak, dan klaster saudara membagi data yang dicakup oleh klaster induknya. Metode ini dapat dikategorikan menjadi dua jenis yaitu, *agglomerative (bottom-up)*, yang setiap objek sebagai klaster tunggal dan secara berulang menggabungkan klaster yang paling mirip hingga seluruh objek tergabung dalam satu klaster besar. Dan *divisive (top-down)*, satu klaster besar yang mencakup semua data, kemudian secara bertahap membagi klaster tersebut hingga mencapai jumlah klaster yang diinginkan (Adek, 2022).

Partitioning Around Medoids (PAM) merupakan salah satu metode dari *clustering*. Metode ini membagi data ke dalam sejumlah cluster dengan memilih *medoid* sebagai representasi setiap *cluster*. *Medoid* adalah titik data aktual dalam cluster yang paling mirip dengan anggota lainnya, sehingga membuat PAM lebih tahan terhadap *outlier*. Metode ini dapat digunakan untuk variabel kuantitatif maupun kualitatif (Sanjari dkk., 2023).

Teknik *clustering* data akan diterapkan dalam sebuah aplikasi dengan algoritma tertentu sehingga tercipta pengolahan data yang efisien. Pemilihan algoritma yang tepat dibutuhkan untuk mendapatkan tingkat keberhasilan dalam penerapannya ke dalam aplikasi. Salah satu metode dalam *clustering* yang umum digunakan adalah *K-Means Clustering*. Algoritma ini bekerja dengan membagi data ke dalam K kelompok, di mana K merupakan jumlah kelompok yang diinginkan. *K-Means Clustering* terkenal cepat dan mampu menangani data dengan jumlah yang besar, sehingga cocok untuk menganalisis data yang besar seperti yang ditemukan pada aplikasi bisnis dan ilmiah (Hendrastuty, 2024).

Penelitian oleh (Meiyanti & Fahada, 2023) dengan judul “*Application of the K-Means Method for Clustering Capture Fisheries Products in North Aceh with A Data Mining Approach*” memaparkan alasan dari pemilihan algoritma *K-Means* dalam penelitiannya. Pemilihan tersebut dilandaskan dari kesederhanaan dan

kemudahan dalam mengimplementasikannya. Disamping itu algoritma ini juga menghasilkan tingkat efisiensi yang tinggi dalam memproses data yang besar. Hal tersebut dapat dilihat dari data yang diolah yaitu hasil tangkapan ikan di Aceh Utara pada tahun 2022. Data tersebut terdiri atas 78 jenis ikan dari 8 kecamatan yang berbeda. Dan hasilnya sebanyak 30 jenis ikan pada data tersebut termasuk ke dalam ikan yang banyak ditangkap dan 48 jenis ikan lainnya termasuk ke dalam ikan yang jarang ditangkap. Maka untuk penelitian ini metode *K-Means Clustering* dipilih.

K-Means Clustering akan digunakan untuk mengelompokkan data pada sebuah perusahaan bernama PT. Amanah Arrazzaq Utama. Menurut direktur utama perusahaan tersebut, perusahaan ini didirikan pada tahun 2022. Perusahaan yang pada awalnya bergerak dibidang transportasi darat dan udara, kini mulai merambat ke penyedia barang dan jasa. Banyaknya bidang yang ingin dirambah sangat membutuhkan analisis data yang tepat untuk menentukan langkah yang tepat ke depannya bagi perusahaan. Pada periode tutup buku 2023, perusahaan ini mencatatkan laba bersih sebesar Rp. 740.000.000 yang didapat dari bidang penyedia jasa pengiriman melalui darat dan udara, serta penyedia barang. Perusahaan ini juga menangani berbagai perusahaan lain sebagai pelanggannya yang diantaranya melayani UMKM hingga perusahaan dengan skala nasional. Dikarenakan data yang besar dengan diversifikasi layanan yang berbeda maka dibutuhkan pengelompokan data hasil proyek yang telah dilakukan oleh perusahaan. Pengelompokan data dilakukan untuk mengevaluasi proyek mana yang memberikan hasil terbaik bagi perusahaan.

Keputusan yang tepat untuk perusahaan sangat dibutuhkan untuk menentukan bagaimana perusahaan tersebut akan berkembang. Maka dibutuhkan pengelolaan data yang tepat sehingga didapatkan informasi yang benar. Informasi tersebut yang akan digunakan oleh perusahaan. Dari permasalahan tersebut maka penulis mengusulkan judul penelitian “Pengelompokan Data Hasil Proyek pada PT. Amanah Arrazzaq Utama dengan Metode *K-Means Clustering* Berbasis Web”. Dengan penelitian tersebut maka diharapkan dapat membantu PT. Amanah Arrazzaq Utama untuk mengevaluasi pemilihan proyek di masa depan dan memilih strategi yang akan digunakan oleh perusahaan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dipaparkan, maka penulis merumuskan masalah yang akan dibahas adalah:

1. Bagaimana membentuk kelompok data berdasarkan *dataset* hasil proyek milik PT. Amanah Arrazzaq Utama menggunakan algoritma *K-Means Clustering* ?
2. Bagaimana mengimplementasikan algoritma *K-Means Clustering* ke dalam sebuah sistem pengelompokan data hasil proyek milik PT. Amanah Arrazzaq Utama berbasis web ?
3. Bagaimana performa *K-Means Clustering* dalam mengelompokan data hasil proyek milik PT. Amanah Arrazzaq Utama ?

1.3 Tujuan Penelitian

Dari pemaparan di atas, penelitian kali memiliki tujuan sebagai berikut:

1. Membentuk kelompok data menggunakan algoritma *K-Means Clustering* berdasarkan *dataset* hasil proyek milik PT. Amanah Arrazzaq Utama.
2. Mengimplementasikan algoritma *K-Means Clustering* ke dalam sistem pengelompokan data berbasis web.

1.4 Manfaat Penelitian

Dari seluruh pemaparan di atas, maka penelitian ini akan menghasilkan manfaat sebagai berikut:

1. Bagi penulis penelitian ini menjadi sarana penerapan pengetahuan yang diperoleh selama menjalani studi.
2. Menjadi masukan bagi PT. Amanah Arrazzaq Utama mengenai pengelompokan hasil proyek kedalam beberapa kategori yaitu, sangat rekomendasi, rekomendasi dan kurang rekomendasi.
3. Meningkatkan pemahaman dan pengetahuan penulis mengenai data *mining* khususnya pengelompokan data menggunakan algoritma *K-Means Clustering* serta pengaplikasiannya ke dalam sebuah sistem aplikasi berbasis *website*.

1.5 Ruang Lingkup dan Batasan Penelitian

Berdasarkan latar belakang dan rumusan masalah di atas, maka ruang lingkup serta batasan penelitian yang penulis berikan adalah sebagai berikut:

1. Penelitian ini hanya menggunakan data hasil proyek milik PT. Amanah Arrazzaq Utama.
2. Data hasil proyek PT. Amanah Arrazzaq Utama berdasarkan laporan keuangan perusahaan, khususnya bagian laporan performa proyek.
3. Data yang digunakan adalah data laporan keuangan dari Juni 2022 sampai dengan Maret 2025 dengan jumlah data sebanyak 153.
4. Metode yang digunakan pada penelitian ini hanya *K-Means Clustering* untuk diimplementasikan ke dalam *website*.
5. Kriteria dari hasil proyek yang digunakan dalam proses pengelompokan adalah tingkat perusahaan lawan transaksi, laba neto, modal, dan jangka waktu.
6. Data akan dikelompokkan berdasarkan tiga kategori yaitu, sangat rekomendasi, rekomendasi, dan kurang rekomendasi.