

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Usaha Mikro, Kecil, dan Menengah (UMKM) memainkan peran penting dalam perekonomian Indonesia, berkontribusi dalam pembentukan banyak bisnis dan penciptaan lapangan kerja. Usaha Mikro Kecil dan Menengah (UMKM) dicirikan sebagai entitas bisnis perorangan atau badan usaha berskala kecil dengan batasan dalam jumlah pekerjaannya, omset, total penjualan, dan total aset. [1].

UMKM biasanya menggunakan bahan baku lokal untuk membuat produk dan layanan bagi masyarakat setempat. Sering dijuluki sebagai ekonomi kerakyatan, UMKM memenuhi kebutuhan sehari-hari masyarakat melalui penawaran mereka[2]. UMKM menyumbang 96,2% dari kesempatan kerja secara nasional dan mewakili porsi terbesar dari semua kategori bisnis. UMKM memainkan peran penting dalam mendorong pertumbuhan dan perkembangan ekonomi Indonesia. [3].

Pada tahun 2023 di Kabupaten Sampang tercatat 36,298 UMKM, yang terbagi menjadi 29,969 usaha mikro, 616 usaha kecil dan 120 usaha menengah menurut yang terdata pada bidang koperasi dan usaha mikro Dinas Koperasi Perindustrian dan Perdagangan Kabupaten Sampang. Sedangkan untuk rekapitulasi usaha mikro di Kecamatan Sampang pada tahun 2023 tercatat 1275 UMKM. Jika dibagi berdasarkan sektor usahanya, sektor makanan dan minuman (mamin) sebanyak 403, sektor perdagangan 373, sektor perikanan 20, sektor peternakan 13, sektor industri 39, sektor jasa 426. UMKM memiliki potensi yang signifikan namun kenyataannya UMKM masih menghadapi berbagai kendala dan hambatan baik diarah internal maupun eksternal yaitu jumlah pertumbuhan UMKM yang sangat besar dan belum pernah dilakukan pengelompokan sehingga Dinas Koperasi Perindustrian dan Perdagangan Kabupaten Sampang kesulitan dalam menentukan

bantuan serta merancang program kerja yang efektif untuk mendukung pengembangan ekonomi lokal [4].

Berdasarkan penjelasan permasalahan UMKM, studi ini secara spesifik meneliti tentang penerapan CRISP-DM dan algoritma *K-Means clustering* yang dapat digunakan sebagai salah satu cara untuk mengelompokkan UMKM di Kecamatan Sampang. Metodologi CRISP-DM terdapat enam tahapan, dimulai dari pemahaman bisnis, pemahaman data, *preprocessing* data, *modelling*, evaluasi, hingga penerapan hasil. Pendekatan ini dapat membantu usaha mikro, kecil dan menengah dalam menyusun dan mengevaluasi data secara efisien untuk mendapatkan wawasan yang berharga untuk membuat pilihan bisnis yang tepat dan strategis, karena pendekatan ini merupakan metodologi serbaguna yang cocok untuk berbagai industri. CRISP-DM dibandingkan metodologi data *mining* yang lain lebih lengkap dan terdokumentasi dengan baik. Mengelompokkan entitas ke dalam kategori-kategori berdasarkan karakteristik yang sama adalah tujuan utama dari *clustering*. Tujuan utama dari pendekatan ini adalah untuk membedakan perbedaan di antara kategori-kategori tersebut[5].

Algoritma *K-Means* menonjol sebagai alat yang penting dalam bidang data *mining* karena kemudahan implementasinya, kecepatan komputasi yang efisien, dan aplikasi yang tersebar luas dalam menangani beragam tantangan komputasi. Selain itu, algoritma ini dikenal sebagai metode pengelompokan yang populer[6]. Pada algoritma ini data akan dibagi menjadi 6 variabel yang akan dikelompokkan ke dalam *cluster* tertentu sesuai dengan karakteristiknya. Data yang diolah adalah data UMKM Kecamatan Sampang tahun 2023. Dataset ini terdiri dari 1275 data dengan atribut seperti nama usaha, jenis usaha, jumlah pekerja, kapasitas produksi, omset, aset dan surat izin. Penelitian yang penulis lakukan ini dengan tujuan yaitu mengimplementasikan metode CRISP-DM menggunakan Algoritma *K-Means clustering* pada UMKM di Kecamatan Sampang untuk mengelompokkan UMKM agar membantu dinas terkait dalam menentukan bantuan serta merancang program kerja yang efektif untuk mendukung pengembangan ekonomi lokal.

Metode evaluasi *elbow* digunakan untuk menentukan jumlah *cluster* optimal dalam algoritma *K-Means*. Kelebihan metode *elbow* adalah mudah dimengerti dan diimplementasikan serta grafik terhadap jumlah *cluster* memberikan pandangan visual yang baik. Metode evaluasi *elbow* memberikan pandangan visual dan sederhana tentang jumlah *cluster* optimal. Selanjutnya untuk mengetahui seberapa baik hasil kualitas *clustering* yaitu menggunakan *Silhouette Coefficient* dan *Davies Bouldin Index* (DBI). *Silhouette Coefficient* adalah metode yang digunakan untuk melihat kualitas serta kekuatan *cluster* dan seberapa baik suatu objek ditempatkan pada suatu *cluster*[7]. Sedangkan DBI adalah metode untuk mengecek hasil *clustering*. Nilai DBI yang semakin mendekati nilai 0 menandakan semakin baik *cluster* yang diperoleh.[8]

Penelitian terhadap proses *clustering* telah banyak dilakukan seperti penelitian yang dilakukan oleh Dwi Astuti, Ade Rahmat Iskandar, Atik Febrianti dalam menentukan strategi promosi UMKM menggunakan metode CRISP-DM dengan algoritma *K-Means Clustering* menghasilkan bahwa *cluster* 1 merupakan *cluster* yang diperhatikan karena memiliki rata-rata aset dan omset yang cukup rendah dibanding dua *cluster* lainnya, berdasarkan data hasil *clustering* yang telah diolah, maka dapat ditentukan beberapa strategi promosi yang dapat dilakukan oleh dinas terkait[9].

Penelitian lain juga dilakukan oleh Ach Muqoddam tentang pengelompokan produksi tambak garam dengan menggunakan metode *cluster K-Means* dan optimasi *cluster* dengan *elbow*, yang menghasilkan nilai K yang paling optimal adalah 3 setelah melakukan uji coba dengan nilai K=7. Nilai SSE untuk K=3 adalah 27.29889, dengan selisih 16.65641, dan grafik *elbow* pada K=3 menunjukkan penurunan yang signifikan dan membentuk siku. Penelitian ini difokuskan pada data tambak garam di Kabupaten Bangkalan yang terdiri dari 141 titik data [10].

Penelitian lain juga dilakukan oleh Yogassetya Suhandi, Ike Kuniati, Siti Norma dalam dalam penerapan metode CRISP-DM dengan Algoritma *K-Means Clustering* untuk segmentasi mahasiswa berdasarkan kualitas akademik yang

menghasilkan pemodelan sistem *dashboard* yang menampilkan hasil *clustering* mahasiswa berdasarkan program studi, presentase nilai mahasiswa berdasarkan perolehan *cluster*, perolehan nilai mahasiswa berdasarkan jenis kelamin berdasarkan masing-masing *cluster* dan informasi IPK mahasiswa berdasarkan *cluster* periode 2009-2018. Hasil penelitian diharapkan dapat digunakan oleh manajemen ITBS untuk mendukung pengambilan keputusan strategi khususnya dibidang akademik[11].

1.2 Perumusan Masalah

1.2.1 Permasalahan

Dari uraian latar belakang yang telah dipaparkan, UMKM di kecamatan Sampang menghadapi beberapa masalah seperti kurangnya modal usaha dan keterbatasan kemampuan sumber daya manusia. Dengan banyaknya jumlah UMKM, diperlukan pengelompokan yang memudahkan Dinas Koperasi, Perindustrian dan Perdagangan Kabupaten Sampang dalam memberikan rekomendasi sesuai kebutuhan setiap kelompok guna mendukung pengembangan ekonomi lokal.

1.2.2 Metode Usulan

Metode yang diusulkan pada penelitian ini yaitu penerapan metode *K-Means Clustering* untuk mengelompokkan UMKM berdasarkan karakteristik yang relevan dengan menggunakan metode optimasi *elbow* untuk mendapatkan hasil *cluster* yang optimal.

1.2.3 Pertanyaan Penelitian

Berdasarkan latar belakang dan rumusan masalah yang telah dijelaskan sebelumnya, maka terbentuklah pertanyaan penelitian sebagai berikut :

1. Bagaimana menentukan pengelompokan UMKM di Kecamatan Sampang menggunakan algoritma *K-Means* dengan optimasi *elbow*?

2. Bagaimana hasil pengelompokan UMKM menggunakan metode *K-Means Clustering* dengan optimasi *elbow*?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan

- a. Untuk menentukan pengelompokan UMKM menggunakan algoritma *K-Means* dengan optimasi *elbow*.
- b. Untuk mengetahui pengelompokan data UMKM menggunakan algoritma *K-Means* dengan optimasi *elbow*.

1.3.2 Manfaat

- a. Dengan mengetahui *cluster* UMKM yang terbentuk dari hasil analisis yang diharapkan dapat membantu pemerintah atau lembaga terkait dalam memberikan rekomendasi sesuai kebutuhan setiap kelompok guna mendukung pengembangan ekonomi lokal.
- b. Dapat mengembangkan pengetahuan analisis *clustering* dengan konsep data mining dengan algoritma *K-Means clustering*.

1.4 Batasan Masalah

Dalam penelitian ini, terdapat batasan-batasan yang perlu diperhatikan, yaitu :

1. Data yang digunakan diperoleh dari Dinas Koperasi, Perindustrian dan Perdagangan Kabupaten Sampang.
2. Data yang digunakan memiliki 6 atribut yaitu Jenis Usaha, Jumlah Tenaga kerja, Kapasitas Produksi, Omset, Aset dan surat izin.
3. Data diolah dengan menggunakan algoritma *K-Means Clustering*.
4. Dataset yang digunakan yaitu 1275 data UMKM kecamatan Sampang.
5. Penentuan jumlah cluster menggunakan metode optimasi *elbow*.

1.5 Sistematika Penulisan

Dalam pembuatan sistematika penulisan penelitian dibuat untuk memudahkan penulis agar menulis dengan runtut serta penulis fokus menganalisis dan terarah, berikut sistematika penulisan laporan penelitian sebagai berikut :

BAB I PENDAHULUAN

Bab ini merupakan latar belakang penelitian, rumusan masalah, permasalahan yang dihadapi, metode yang diusulkan, pertanyaan penelitian, tujuan dan manfaat penelitian, batasan masalah yang dibahas, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini menjelaskan mengenai kumpulan teori ilmiah yang melandasi masalah penelitian yang terdiri UMKM, Data *Mining*, *Clustering*, Algoritma *K-Means* dan penelitian terkait.

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan dengan rinci mengenai perancangan sistem yang akan dibuat saat penelitian seperti alur penelitian, dataset, tempat, dan waktu penelitian, algoritma sistem pengelompokan, dan pengujian.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini menjelaskan mengenai hasil dan pembahasan dari penelitian yang telah dilakukan.

BAB V PENUTUP

Pada bab ini menjelaskan mengenai hasil kesimpulan dan saran dari penelitian

PUSTAKA Berisi daftar referensi dari penelitian yang telah dilakukan.