

BAB I

PENDAHULUAN

1.1. Latar Belakang

UMKM merupakan singkatan dari Usaha Mikro, Kecil, Menengah yang pada dasarnya umkm dikendalikan oleh perorangan, kelompok, usaha kecil atau rumah tangga (Ramba & Rantung, 2023). Saat ini terdapat sekitar 64,2 juta unit usaha yang tergolong dalam UMKM dan mampu berkontribusi terhadap Produk Domestik Bruto (PDB) sebesar 61,7% yang setara Rp. 8.573,89 trilyun. Selain itu, UMKM juga mampu menyerap tenaga kerja sekitar 97% dari total tenaga kerja (Sudrajat et al., 2022). Kontribusi tersebut menjadikan UMKM sebagai salah satu pendorong utama dalam pertumbuhan perekonomian di Indonesia. Peran UMKM tidak hanya terbatas pada aspek ekonomi, tetapi juga mencakup aspek sosial, seperti pemerataan pendapatan, pengurangan tingkat pengangguran, serta penguatan struktur ekonomi berbasis kerakyatan (Ramdhan & Erlangga, 2025). Pada tingkat daerah, seperti Kabupaten Sidoarjo di provinsi Jawa Timur, UMKM berkembang secara signifikan dan menjadi motor penggerak utama ekonomi lokal dengan angka pertumbuhan ekonomi sebesar 5,54% melampaui capaian nasional dan Jawa Timur sebesar 5,03% dan 4,93% (Ismail, 2025). Meskipun demikian, dinamika perkembangan UMKM di daerah masih menghadapi berbagai tantangan, terutama terkait keterbatasan data yang terintegrasi, segmentasi UMKM yang belum jelas, serta minimnya intervensi kebijakan yang didasarkan pada analisis data yang mendalam (Ramdhan & Erlangga, 2025).

Salah satu pendekatan yang dapat digunakan untuk mengatasi tantangan tersebut adalah melakukan segmentasi atau pengelompokan UMKM berdasarkan kemiripan karakteristik. Pendekatan ini menjadi semakin relevan seiring dengan berkembangnya teknologi pengolahan data dan analisis berbasis kecerdasan buatan, seperti *machine learning* dan *data mining*. Perkembangan teknologi ini memungkinkan proses klasterisasi dilakukan secara lebih akurat, cepat, dan mampu menangani volume data UMKM yang besar serta heterogen (Gordon & Unus, 2021). Penggunaan teknologi tersebut tidak hanya mempermudah identifikasi pola dan kecenderungan dalam data UMKM, tetapi juga memberikan landasan bagi

praktik pengambilan keputusan berbasis data (*data-driven decision making*) yang lebih tepat sasaran dalam penyusunan kebijakan pengembangan UMKM ke depan. Dalam implementasinya, pengelompokan UMKM dihadapkan pada karakteristik data yang heterogen dan kompleks. Metode K-Means merupakan salah satu metode clustering yang telah banyak digunakan dalam berbagai penelitian. Namun, karena sifatnya yang sudah umum digunakan, diperlukan eksplorasi metode lain yang dapat memberikan sudut pandang berbeda dalam proses pengelompokan data, seperti pendekatan berbasis fuzzy.

Salah satu algoritma *clustering* adalah Fuzzy C-Means (FCM) yang diperkenalkan oleh Bezdek (1981). Keunggulan utama FCM terletak pada kemampuannya dalam menangani ketidakpastian data melalui pendekatan *soft clustering*, di mana setiap data dapat menjadi anggota dari lebih dari satu kluster dengan derajat keanggotaan tertentu (Fawas & Handhayani, 2025). Pendekatan tersebut berbeda dengan metode *hard clustering* seperti *K-Means*, yang mengelompokkan setiap data hanya ke dalam satu kluster secara mutlak. Berbagai studi telah menunjukkan bahwa Fuzzy C-Means telah banyak digunakan dalam analisis data sosial-ekonomi termasuk pemetaan kelompok UMKM, karena kemampuannya menangani data yang tidak memiliki batas kluster tegas. Misalnya, penelitian yang dilakukan oleh Rouza & Fimawa (2020) menerapkan FCM untuk pengelompokan UMKM di Kabupaten Rokan Hulu mendapatkan nilai validasi rata-rata mendekati 1 yang menunjukkan bahwa FCM memiliki tingkat akurasi yang tinggi sebesar 80%-90%. Hal tersebut menunjukkan bahwa algoritma ini mampu memberikan representasi kluster yang lebih realistis dalam konteks data UMKM yang heterogen serta metode ini lebih adaptif terhadap variasi karakteristik usaha (Rouza & Fimawahib, 2020). Temuan tersebut menguatkan bahwa FCM merupakan metode yang tepat digunakan pada data UMKM, yang umumnya penuh ketidakpastian pelaporan, dan tumpang tindih karakteristik antar pelaku. Namun demikian, FCM konvensional atau Fuzzy C-Means Type-1 memiliki sejumlah keterbatasan mendasar ketika diterapkan pada data yang kompleks, tidak homogen, atau mengandung *noise*. Hal ini disebabkan karena FCM Type-1 menggunakan fungsi keanggotaan tunggal dengan parameter fuzzifier tetap, yang mengatur tingkat “kabur” antar kluster.

Hal ini disebabkan karena FCM Type-1 menggunakan fungsi keanggotaan tunggal dengan parameter fuzzifier tetap yang mengatur tingkat “kabur” antar kluster. Dalam banyak kasus, satu nilai fuzzifier yang digunakan secara global tidak mampu menyesuaikan dengan variasi distribusi maupun kepadatan data, sehingga hasil klusterisasi menjadi tidak stabil dan akurasi menurun. Selain itu, pendekatan ini juga tidak mampu merepresentasikan ketidakpastian yang timbul akibat variabilitas data, noise pengukuran, maupun tumpang tindih antar kluster (Yin et al., 2022). Untuk mengatasi kelemahan tersebut, dikembangkan metode Fuzzy C-Means Type-2 (FCM Type-2) yang mampu menangani ketidakpastian melalui konsep upper dan lower membership function. Penelitian terbaru menunjukkan bahwa Fuzzy Type-2 lebih unggul dalam data yang mengandung variabilitas tinggi, noise, dan ambiguitas (Mishro et al., 2021).

Selain itu, penentuan jumlah kluster yang sesuai dengan data sangat penting agar hasil *Clustering* menggunakan algoritma Fuzzy C-Means Type-2 menjadi optimal. Salah satu metode yang relevan digunakan dalam mengevaluasi jumlah kluster pada penelitian fuzzy adalah *Modified Partition Coefficient* (MPC) yang merupakan penyempurnaan dari Partition Coefficient karena mampu menghilangkan bias terhadap jumlah kluster. MPC memberikan ukuran yang lebih objektif dalam menilai tingkat kejelasan partisi fuzzy. Penelitian yang dilakukan oleh Nida et al. (2024) menunjukkan bahwa *Modified Partition Coefficient* dapat digunakan secara efektif dalam menilai kualitas hasil kluster. Jumlah kluster optimal dalam penelitian tersebut adalah sebanyak 2 kluster dengan skor MPC sebesar 0,6506 yang mengindikasikan tingkat kejelasan partisi fuzzy dengan baik (Nida et al., 2024). Dengan skor tersebut menunjukkan bahwa penggunaan MPC dianggap mampu memberikan gambaran yang representatif mengenai kualitas kluster dalam data yang mengandung ketidakpastian dan tumpang tindih antar kelompok.

Oleh karena itu, penelitian ini bertujuan untuk mengevaluasi hasil clustering menggunakan MPC pada algoritma *Fuzzy C-Means Type-2* dalam pemetaan UMKM di Kabupaten Sidoarjo. Evaluasi ini dilakukan untuk menilai kualitas pengelompokan yang dihasilkan serta menentukan jumlah cluster yang optimal berdasarkan struktur data yang tersedia. Penelitian ini juga bertujuan untuk mengidentifikasi pola karakteristik UMKM secara lebih mendalam melalui

pendekatan Fuzzy C-Means Type-2, dengan fokus pada peningkatan kualitas hasil Clustering. Dengan pemetaan yang tepat sasaran, diharapkan dapat membantu pemerintah dan pelaku usaha dalam merancang strategi pengembangan yang sesuai dengan kebutuhan dan potensi masing-masing kelompok UMKM. Pendekatan ini diharapkan mampu memberikan solusi yang lebih tepat dalam pengembangan UMKM di Kabupaten Sidoarjo serta mendukung pengambilan keputusan berbasis data untuk kebijakan dan strategi yang lebih efektif dan efisien.

1.2. Perumusan Masalah

Pada bagian perumusan masalah, diuraikan mengenai permasalahan, metode usulan serta pertanyaan penelitian yang diuraikan sebagai berikut:

1.2.1. Permasalahan (Problem Statement)

Permasalahan utama dalam penelitian ini adalah belum tersedianya pemetaan UMKM Kabupaten Sidoarjo yang akurat, stabil, dan representatif akibat kompleksitas karakteristik data UMKM yang bersifat homogen serta mengandung ketidakpastian. Data UMKM terdiri atas variabel numerik seperti modal, omzet, dan jumlah tenaga kerja, serta variabel kategorikal seperti jenis usaha dan status perizinan, dan platform digital yang sering kali disertai fluktuasi nilai dan ketidakkonsistenan pencatatan. Kondisi tersebut membuat metode klasterisasi konvensional, seperti K-Means maupun FCM Type-1, kurang mampu memberikan hasil klaster yang tepat karena bersifat deterministik, sensitif terhadap outlier, dan tidak dapat menangani ambiguitas serta tumpang tindih antar klaster. Selain itu, metode FCM Type-1 hanya menggunakan satu fungsi keanggotaan dengan parameter fuzzifier tunggal sehingga tidak dapat sepenuhnya memodelkan variasi dan noise yang muncul pada data UMKM. Hal ini menyebabkan hasil segmentasi menjadi kurang stabil serta tidak mencerminkan kondisi nyata di lapangan. Dengan demikian, diperlukan metode yang dapat menangkap ketidakpastian data melalui pendekatan yang lebih fleksibel dan adaptif. *Fuzzy C-Means Type-2* menjadi solusi potensial karena bekerja berdasarkan dua tingkat keanggotaan (*upper-lower membership*) yang mampu memodelkan *Footprint of Uncertainty* (FOU), sehingga lebih cocok untuk data UMKM yang tidak pasti, dinamis, dan tumpang tindih.

Oleh karena itu, permasalahan inti dalam penelitian ini berkaitan dengan

kebutuhan untuk merancang metode klasterisasi yang mampu mengakomodasi karakteristik data UMKM secara komprehensif dan menghasilkan segmentasi yang lebih stabil, akurat, dan dapat digunakan sebagai dasar perencanaan kebijakan pengembangan UMKM yang tepat sasaran

1.2.2. Metode Usulan

Metode yang diusulkan dalam penelitian ini adalah *Fuzzy C-Means Type-2* (FCM Type-2) sebagai pendekatan klasterisasi yang mampu menangani ketidakpastian dan variabilitas tinggi pada data UMKM. FCM Type-2 menggunakan konsep *upper membership function* dan *lower membership function* untuk menghasilkan pemodelan klaster yang lebih adaptif dan tidak bergantung pada satu nilai parameter fuzzifier saja. Metode ini dinilai lebih unggul pada data yang heterogen, mengandung noise, dan memiliki batas klaster yang tidak tegas. Selain itu, dalam konteks evaluasi kualitas hasil klasterisasi fuzzy indeks validitas klaster menjadi aspek penting untuk mengukur kebaikan partisi yang dihasilkan. Salah satu indeks validasi klaster tersebut adalah *Modified Partition Coefficient* (MPC) yang merupakan modifikasi dari *partition coefficient* yang memiliki kekurangan dalam hal kecenderungan menghasilkan nilai yang meningkat seiring bertambahnya jumlah klaster sehingga menimbulkan bias dan sulit digunakan untuk membandingkan performa klasterisasi pada variasi jumlah klaster yang berbeda. MPC memperbaiki kelemahan tersebut dengan melakukan normalisasi terhadap nilai PC sehingga menghasilkan rentang nilai yang konsisten antara 0 hingga 1 dan mampu memberikan penilaian yang lebih objektif, stabil, dan adil dalam mengevaluasi kualitas partisi fuzzy.

Pendekatan ini diharapkan dapat menghasilkan segmentasi UMKM yang lebih konsisten, dapat diinterpretasikan dengan baik, serta relevan untuk mendukung penyusunan strategi pembinaan, bantuan modal, hingga program digitalisasi UMKM di Kabupaten Sidoarjo.

1.2.3 Pertanyaan Penelitian

Berdasarkan uraian latar belakang, dapat diambil pernyataan penelitian diantaranya sebagai berikut:

1. Bagaimana hasil penerapan algoritma *Fuzzy C-Means Type-2* dengan optimasi *Modified Partition Coefficient*?
2. Bagaimana hasil pemetaan UMKM Kab. Sidoarjo menggunakan *Fuzzy C-Means Type-2* dengan optimasi *Modified Partition Coefficient*?

1.3. Tujuan & Manfaat

Pada bagian tujuan dan manfaat, diuraikan sasaran penelitian serta kontribusi yang diharapkan sebagai berikut:

1.3.1 Tujuan Penelitian

1. Untuk mengetahui jumlah kluster paling optimal pada data UMKM Sidoarjo sehingga diperoleh pengelompokan yang memiliki tingkat kemiripan tinggi dalam kluster dan perbedaan yang jelas antar kluster.
2. Menerapkan algoritma *Fuzzy C-Means Type-2* dengan parameter hasil optimasi MPC guna menghasilkan kluster UMKM yang lebih stabil terhadap ketidakpastian data dan memiliki kualitas lebih baik dibandingkan metode standar
3. Untuk mengetahui hasil pemetaan UMKM di Kabupaten Sidoarjo menggunakan algoritma *Fuzzy C-Means Type-2* dengan optimasi *Modified Partition Coefficient*, sehingga dapat diperoleh struktur kluster yang terbentuk berdasarkan karakteristik data.

1.3.2 Manfaat

Berdasarkan rumusan masalah yang ada, dapat ditentukan manfaat dari penelitian ini diantaranya adalah sebagai berikut:

1. Studi ini dapat mengembangkan wawasan mengenai jumlah kluster optimal data UMKM yang didasarkan pada nilai index tertinggi, sehingga akan terbentuk sebuah kluster atau kelompok yang berkualitas dan sesuai dengan karakteristik data.
2. Memberikan solusi berbasis data yang dapat diimplementasikan oleh pemerintah daerah dan pelaku usaha untuk merancang strategi pengembangan yang lebih efektif dan efisien sesuai dengan potensi dan kebutuhan masing-masing kelompok UMKM.

3. Memberikan gambaran hasil pemetaan umkm sebagai bahan pertimbangan pengambilan keputusan berbasis data.

1.4 Batasan-batasan

Berikut batasan-batasan pada penelitian ini:

1. Data yang digunakan ialah data UMKM Kabupaten Sidoarjo yang diperoleh dari Dinas Koperasi dan Usaha Mikro Kabupaten Sidoarjo.
2. Data diolah menggunakan algoritma *Fuzzy C-Means Type-2* untuk klusterisasi.
3. Data yang digunakan memiliki 5 atribut yakni Modal, Omset perbulan, Tenaga kerja, Izin usaha, PlatformDigital.
4. Jumlah kluster dioptimalkan menggunakan metode *Modified Partition Coefficient*.

1.5 Sistematika Proposal

Sistematika penulisan proposal ini penulis susun dengan mengikuti aturan penulisan skripsi pada program studi S1 Sistem Informasi Universitas Trunodjoyo Madura. Berikut merupakan metodologi penulisan proposal pada penelitian ini:

BAB I PENDAHULUAN

Bagian ini menjelaskan latar belakang penelitian, perumusan masalah, pertanyaan penelitian, tujuan dan manfaat penelitian, batasan penelitian, serta sistematika penulisan proposal.

BAB II TINJAUAN PUSTAKA

Bagian ini mencakup kajian literatur terkait konsep UMKM, Data Mining, *Clustering*, CRISP-DM, Himpunan Fuzzy, Fuzzy C-Means Type-2, *Modified Partition Coefficient* serta penelitian terdahulu sebagai dasar teoritis penelitian.

BAB III METODOLOGI PENELITIAN

Bagian ini menjelaskan metode yang digunakan dalam penelitian, meliputi metodologi penelitian, variabel penelitian, tahapan CRISP-DM hingga skenario pengujian yang akan dilakukan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian menggunakan algoritma *fuzzy c-means type-2*, analisa data, dan evaluasi menggunakan *Modified Partition Coefficient*, serta pembahasan terkait penemuan penelitian.

BAB V PENUTUP

Bab ini menghasilkan kesimpulan dan saran yang dapat diterapkan.

REFRENSI Bab berisi tentang keterangan sumber rujukan yang digunakan dalam penyusunan Tugas Akhir. Penyusunan referensi pada penelitian ini dilakukan secara otomatis menggunakan aplikasi *mendeley*.