

BAB I

PENDAHULUAN

1.1. Latar Belakang

Masalah kemiskinan merupakan hal yang sangat kompleks. Di wilayah Kecamatan Bantul, seorang warga disebut sebagai keluarga miskin berdasarkan beberapa aspek seperti aspek pangan, sandang, papan, penghasilan, kesehatan, pendidikan, kekayaan, air bersih, listrik maupun jumlah jiwa. Dalam manajemen penanggulangan dan pengentasan kemiskinan di Kabupaten Bantul telah ditetapkan satu data keluarga miskin dengan berdasarkan pada Peraturan Bupati No. : 21 A Tahun 2007 tgl 1 April 2007 tentang Indikator Keluarga Miskin Kabupaten Bantul. Untuk mendapatkan data kemiskinan secara sistematis dan rutin maka harus terdapat jaringan pelaksana yang kuat secara berjenjang dari tingkat Kabupaten sampai ke tingkat Pedukuhan.

Sumber data dari BPS menunjukkan bahwa di Indonesia mempunyai angka kemiskinan pada bulan September 2014 relatif masih tinggi yaitu sekitar 27,72 juta orang (10,96%), angka ini berkurang 0,552 juta orang dari data bulan Maret 2014 yaitu sebesar 28,28 juta orang. Jumlah penduduk miskin di provinsi Daerah Istimewa Yogyakarta tergolong tinggi, sejumlah 532.590 penduduk dengan prosentase kemiskinan 14,55%. Berdasarkan angka tersebut terlihat masih tingginya angka kemiskinan yang ada di wilayah Yogyakarta secara umum.

BKKBN (Badan Kependudukan dan Keluarga Berencana Nasional) sebagai badan yang bertugas menghimpun data statistik kemiskinan di Kabupaten Bantul merasa kesulitan dalam pendistribusian berbagai macam bantuan yang ada karena data yang

sulit diperoleh dan tingkat kemiskinan warga yang susah diukur secara pasti. BKKBN berharap terdapat semacam kelas-kelas kemiskinan menurut kondisi keluarga yang disajikan dalam bentuk visualisasi pemetaan. Dengan pemetaan ini harapannya BKKBN dapat mempermudah dalam melihat distribusi daerah-daerah miskin di Kecamatan Bantul.

Upaya-upaya untuk membantu program pengentasan kemiskinan di daerah Bantul pernah dilakukan melalui penelitian oleh Rianto (2008) dan Ernawati (2012) tetapi penelitiannya belum mampu menunjukkan visualisasi penyebaran keluarga miskin. Rianto melakukan klasifikasi keluarga miskin menggunakan *Analytical Hierarchy Process* (AHP) tetapi belum memetakan hasil klasifikasinya. Berbeda dengan Rianto, Ernawati (2012) menggunakan metode deskriptif kuantitatif dan kualitatif untuk memetakan potensi penduduk miskin per kecamatan di Kabupaten Bantul. Metode kuantitatif ini digunakan untuk skoring (pembobotan) terhadap variabel-variabel penentu potensi rumah tangga miskin dan variabel-variabel penentu potensi wilayah. Kelemahan dari penelitian ini adalah pemetaan potensi keluarga miskin disajikan per kecamatan sehingga masih terlalu umum.

Berdasarkan pada kondisi tersebut, maka dalam proposal ini akan diusulkan model *clustering* untuk mendapatkan klaster-klaster kemiskinan dengan menganalisa aspek yang berpengaruh maupun tidak. Upaya tersebut dilakukan melalui pembuatan suatu alat bantu berupa aplikasi dengan menggunakan metode *Fuzzy C-Means* (FCM) untuk mengetahui pola penduduk miskin dengan seleksi fitur *Information Gain*. Hasil *clustering* akan divisualisasikan dalam bentuk pemetaan. Visualisasi hasil *clustering*

akan disajikan pada area antar pedukuhan. Visualisasi dilakukan dengan tujuan agar mempermudah dalam melihat lokasi-lokasi penyebaran keluarga miskin.

Metode FCM dipilih karena suatu warga mungkin dapat menjadi anggota dari masing-masing cluster (kelompok) dengan derajat keanggotaan yang berbeda antara 0 dan 1. Pada penelitian yang dilakukan oleh Wang (2004) dinyatakan bahwa Algoritma *Fuzzy C-Means* sangat bergantung pada pemilihan matriks awal untuk proses klasterisasi. Dimana Algoritma *Fuzzy C-Means* juga bergantung pada fitur bobot yang mempengaruhi jarak antar klaster yang terbentuk. Sehingga penelitian yang dilakukan tersebut dinyatakan perlu adanya suatu penyesuaian fitur bobot pada Algoritma *Fuzzy C-Means*, hal ini dikuatkan oleh penelitian yang dilakukan oleh Ingunn (2008). Dengan menggunakan metode ini, sebuah keluarga bisa jadi masuk lebih dari satu klaster. Sebagai contoh, sebuah keluarga berdasarkan aspek yang digunakan, dapat masuk dalam 2 klaster sekaligus, misalnya klaster miskin dan rawan miskin dengan nilai keanggotaan antara 0 – 1.

Metode *Information Gain* digunakan sebagai salah satu cara untuk mengurangi dimensi dengan memilih fitur yang tepat atau hanya menggunakan atribut-atribut yang diperlukan. Secara konseptual, pemilihan sub-set fitur merupakan suatu proses pencarian terhadap semua kemungkinan sub-set fitur. *Information Gain* dipilih untuk menyeleksi fitur karena proses seleksi fitur akan dilakukan sebelum proses *clustering* dan data yang digunakan berupa data kategorial sehingga cocok menggunakan metode ini.

1.2. Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan di atas, maka rumusan masalah penelitian ini adalah bagaimana mengembangkan suatu sistem yang mampu menangani proses clustering data penduduk miskin untuk mencari pola cluster kemiskinan suatu kecamatan dengan aspek-aspek yang sesuai dimana hasilnya akan divisualisasikan melalui pemetaan pada area antar pedukuhan

1.3. Batasan Masalah

Permasalahan yang akan dibahas pada penelitian ini memiliki ruang lingkup yang cukup luas, sehingga perlu dilakukan pembatasan sebagai berikut :

1. Data kemiskinan yang akan dianalisa adalah data statistik daerah khusus keluarga miskin yang dihimpun oleh BKKBN Bantul
2. Clustering dilakukan hanya untuk penduduk Kecamatan Bantul
3. Pelabelan pada peta diwakili oleh pedukuhan

1.4. Tujuan

Tujuan yang ingin dicapai dalam penelitian ini adalah mengembangkan suatu sistem dengan menggunakan teknik data mining yang dapat digunakan untuk penentuan cluster penduduk miskin menggunakan Fuzzy C-Means dengan seleksi atribut information gain berdasarkan indikator yang telah ditetapkan BKKBN beserta visualisasi dalam bentuk pemetaan pada area antar pedukuhan.

1.5. Manfaat

Adapun manfaat dari penelitian ini diantaranya adalah :

1. Dapat membantu BKKBN untuk mengidentifikasi dan mengelompokkan penduduk miskin berdasarkan kondisi keluarganya.
2. Metode yang digunakan pada penelitian ini dapat digunakan pada riset mengenai analisa data yang lain, dimana keakuratan metode ini akan coba dibandingkan dengan metode lainnya.