

# **BAB I**

## **PENDAHULUAN**

### **1.1 LATAR BELAKANG MASALAH**

Kemiskinan merupakan salah satu persoalan mendasar yang menjadi pusat perhatian pemerintah di negara manapun. Menurut Kunarjo dalam Badrul Munir (2002:10), suatu negara dikatakan miskin biasanya ditandai dengan tingkat pendapatan perkapita rendah, mempunyai tingkat pertumbuhan penduduk yang tinggi (lebih dari 2 persen per tahun). Ukuran atau kategori kemiskinan menurut BPS (2005) antara lain penduduk miskin dikatakan sangat miskin apabila kemampuan untuk memenuhi konsumsi makanan hanya mencapai 1900 kalori per orang per hari ditambah kebutuhan dasar non makanan, atau setara dengan Rp 120.000,- per orang atau perbulan, penduduk dikatakan miskin apabila kemampuan memenuhi konsumsi makanan hanya mencapai antara 1900-2100 kalori per orang ditambah kebutuhan dasar non makanan, atau setara Rp.150.000,- per orang per bulan, dan penduduk dikatakan mendekati miskin apabila kemampuan memenuhi konsumsi hanya mencapai antara 2100-2300 kalori ditambah kebutuhan dasar non-makanan atau setara Rp 175.000,- per orang per hari. Mengacu pada permasalahan tersebut maka diperlukan sebuah metode yang dapat digunakan untuk pengklasteran data penduduk miskin yaitu dengan menerapkan teknik *data mining*.

*Data Mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan

mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Tujuan utama *data mining* adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki, Mujib Ridwan et al. (2013 : 60). Salah satu metode *data mining* yang dapat menyelesaikan permasalahan diatas adalah *Clustering* seperti penelitian yang dilakukan oleh Zainul Aras Z dan Sarjono (2016) yang menggunakan algoritma *K-Means* untuk menentukan prioritas penerima bantuan bedah rumah, metode *Clustering K-Means* digunakan agar prioritas penduduk tidak mampu bisa lebih berkualitas dan efektif. Rizky Dwi Setiawan dkk (2016) melakukan perhitungan pemetaan tingkat kemiskinan menggunakan algoritma *K-Means Clustering*, adapun rentang dari tingkat kesejahteraan memperoleh 3 kelompok, kelompok pertama yaitu hampir miskin dengan rentang nilai  $<2.117$ , kelompok kedua yaitu miskin dengan rentang nilai  $\geq 2.117 - <2.807$ , dan kelompok ketiga sangat miskin  $\geq 2.807$ .

Hal inilah yang melatar belakangi penulis untuk melakukan penelitian guna memberi solusi terhadap masalah yang terjadi dengan mengangkat judul “PENERAPAN *DATA MINING* UNTUK *CLUSTERING* DATA PENDUDUK MISKIN PROVINSI INDONESIA BERDASARKAN WILAYAH MENGGUNAKAN ALGORITMA *K-MEANS*”.

## 1.2 RUMUSAN MASALAH

Berdasarkan latar belakang masalah diatas maka dapat dirumuskan permasalahan yang akan diteliti yaitu:

1. Bagaimana penerapan *Data Mining* untuk *Clustering* data penduduk miskin Provinsi Indonesia berdasarkan wilayah menggunakan algoritma *K-means*?”.
2. Bagaimana mengevaluasi dan menganalisis dari hasil penerapan *Clustering* dengan Algoritma *K-Means*?”.

## 1.3 BATASAN MASALAH

Untuk lebih memfokuskan pada permasalahan yang akan diteliti, maka penelitian ini dibatasi sebagai berikut:

1. Penelitian ini menggunakan metode *Clustering* dengan algoritma *K-Means*.
2. Data yang digunakan merupakan data yang diambil dari dataset.
3. Alat bantu yang digunakan pada penelitian ini adalah aplikasi *WEKA*.
4. Atribut yang digunakan adalah : Persentase Penduduk Miskin, Persentase Penduduk Tidak Tamat Sekolah Dasar, Persentase Penduduk Tamat Sekolah Dasar dan Sekolah Menengah Pertama, Persentase Penduduk Tamat Sekolah Menengah Atas dan Indeks Pembangunan Manusia

## **1.4 TUJUAN DAN MANFAAT PENELITIAN**

### **1.4.1 Tujuan Penelitian**

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Menerapkan *Clustering* data penduduk miskin Berdasarkan Wilayah dengan menggunakan algoritma *K-Means*.
2. Mengevaluasi dan menganalisis hasil perhitungan *Clustering* dengan algoritma *K-Means* pada penduduk miskin Provinsi Indonesia berdasarkan wilayah.

### **1.4.2 Manfaat Penelitian**

Manfaat dari penelitian ini yaitu :

1. Dapat mengetahui prioritas wilayah penduduk miskin.
2. Penulis dapat menambah ilmu dan wawasan baru mengenai penerapan data mining untuk *Clustering* data penduduk miskin berdasarkan wilayah menggunakan algoritma *K-Means*.
3. Dapat digunakan sebagai acuan untuk penelitian berikutnya.

## **1.5 SISTEMATIKA PENULISAN**

Adapun sistematika penulisan tugas akhir ini, penulis menguraikan dalam beberapa bab yaitu:

### **BAB I : PENDAHULUAN**

Pada bab ini akan diuraikan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat penulisan, dan sistematika penulisan.

**BAB II : LANDASAN TEORI**

Pada bab landasan teori ini membahas tentang teori-teori dan pendapat para ahli yang berhubungan dengan permasalahan yang di analisis. Teori-teori yang digunakan antara lain mengenai penerapan *Data Mining* untuk *Clustering* data penduduk miskin, algoritma *K-Means* dan *WEKA*.

**BAB III : METODOLOGI PENELITIAN**

Pada bab ini menjelaskan tentang kerangka kerja penelitian, metode pengumpulan data, metode *Clustering*, serta alat bantu yang digunakan dalam penelitian ini.

**BAB IV : ANALISIS PROSES PERHITUNGAN**

Pada bab ini dilakukan perhitungan analisis menggunakan metode *K-Means Clustering* terhadap data-data penduduk miskin yang tersedia.

**BAB V : HASIL ANALISIS DAN VISUALISASI**

Pada bab ini akan ditampilkan hasil dari analisis dan bentuk visualisasi analisis dari *tools WEKA* yang digunakan.

**BAB VI : PENUTUP**

Bab ini yang berisikan kesimpulan-kesimpulan yang di ambil dari analisis serta saran-saran yang mencakup keseluruhan dari hasil penelitian.