

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kumpulan data yang terdapat pada suatu institusi baik yang bersifat komersial seperti perusahaan maupun yang bersifat nirlaba seperti organisasi pemerintah merupakan salah satu aset yang dapat digali informasinya. Data tersebut umumnya berjumlah besar dan tersimpan dalam suatu basis data. Salah satu teknik yang digunakan untuk penggalian informasi yaitu dengan *data mining*. *Data mining* merupakan proses mengekstrasi informasi dan pola yang berguna dari jumlah data yang besar, disebut juga sebagai *knowledge discovery process* [19].

Salah satu aset yang dapat digali adalah data dari perusahaan perbankan Islam yang menjadi sorotan dunia saat ini. Krisis global baru-baru ini membawa keuangan Islam maju sebagai alternatif dalam hal investasi dan perbankan [21]. Perbankan Islam telah menjadi bagian integral dari struktur keuangan global terutama dengan kekebalan terhadap krisis perbankan dan keuangan baru-baru ini [1].

Bank Pembiayaan Rakyat Syariah (BPRS) merupakan salah satu bank syariah yang menjalankan kegiatan usahanya dengan prinsip syariah yang kegiatannya tidak memberikan jasa dalam lalu lintas pembayaran. Berdasar-

kan laporan publikasi Lembaga Penjamin Simpanan (LPS), jumlah BPRS di Indonesia yang terdaftar hingga 2016 lalu adalah 164 bank. Dari 164 BPRS tersebut sebanyak 3 BPRS sudah dilikuidasi dan sedang dalam proses likuidasi dari tahun 2013-2016.

Berbeda jauh dengan Bank Perkreditan Rakyat yaitu sebanyak 26 sudah dilikuidasi atau sedang dalam proses likuidasi. Oleh sebab itu perlu adanya analisis lebih lanjut, agar dapat diketahui kelompok bank mana saja yang perlu diberikan pengawasan khusus agar tidak terjadi penambahan penutupan bank khususnya BPRS.

Salah satu teknik *data mining* yang digunakan untuk mengelompokkan data adalah dengan *cluster analysis*. *Cluster analysis* merupakan proses mempartisi sekumpulan objek data atau observasi ke dalam himpunan bagian. Objek dalam kelompok memiliki kemiripan satu sama lain, namun berbeda dengan objek dalam kelompok lain [12]. Pada umumnya, *clustering* untuk data kategorik dan numerik tidak dapat digunakan dalam algoritma *cluster* yang sama. Permasalahan yang sering di dapat dalam analisis *cluster* adalah jenis data yang berskala campuran kategorik dan numerik.

Kerap pula dilakukan teknik transformasi data pada data campuran yakni dengan mentransformasi data numerik menjadi data kategorik atau sebaliknya. Metode transformasi memiliki kelebihan dapat mengurangi kompleksitas dalam komputasi. Akan tetapi, metode tersebut memiliki kelemahan dalam memilih transformasi yang sesuai agar tidak kehilangan banyak informasi dari keaslian datanya.

Selain *clustering* dengan metode transformasi, He, Xu dan Deng mengembangkan sebuah metode *clustering* untuk data campuran yakni metode *ensemble* [10]. *Clustering ensemble* adalah teknik *clustering* untuk menggabungkan hasil *clustering* beberapa algoritma untuk mendapatkan *cluster* yang lebih baik.

Menurut Zengyou He metode *ensemble* untuk data campuran adalah dengan *Algoritma Cluster Ensemble Based Mixed Data Clustering* (AlgCEB-MDC) [11]. Tahapnya yakni *clustering* secara parsial data kategorik dan numerik dengan masing-masing algoritma yang dipilih. Kemudian dari hasil *clustering* tersebut digabungkan untuk selanjutnya dikelompokkan dengan algoritma yang dipilih. Berdasarkan hasil penelitian yang dilakukan Alvionita yang membandingkan metode *ensemble* ROCK dan metode *Similarity Wirght and Filter Method* (SWFM) menyatakan bahwa kinerja metode *ensemble* ROCK lebih baik dari pada metode SWFM [2]

Pada penelitian ini, dilakukan *clustering* dengan metode *ensemble* dalam mengelompokkan BPRS dengan penggabungan *algoritma* ROCK (*Robust Clustering using linKs*) untuk data kategorik, dan *Algoritma Hierarchical Agglomerative Nesting* (AGNES) untuk data numerik. Metode *ensemble* ROCK merupakan penggabungan hasil *cluster* terbaik dari data kategorik dan hasil *cluster* terbaik dari data numerik dan dipandang sebagai data baru dengan tipe kategorik, kemudian dilakukan *clustering* dengan menggunakan metode ROCK untuk mendapatkan hasil akhir.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut maka yang menjadi masalah pada penelitian ini adalah bagaimana *cluster ensemble* ini dapat digunakan untuk *clustering* objek dan bagaimana ilustrasinya pada data.

## 1.3 Batasan Masalah

Pada penelitian ini analisis *cluster ensemble* diilustrasikan untuk *clustering* BPRS, berdasarkan data yang diperoleh dari Otoritas Jasa Keuangan (OJK). *Clustering* dilakukan untuk mengelompokkan objek ke dalam 2 - 5 *cluster*.

## 1.4 Tujuan Penelitian

Adapun tujuan penelitian ini adalah:

1. Untuk mengetahui hasil *cluster* yang terbentuk menggunakan metode *ensemble* ROCK untuk data berskala campuran kategorik dan numerik.
2. Untuk mengetahui karakteristik dari hasil *cluster* yang terbentuk menggunakan metode *ensemble* ROCK.

## 1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah:

1. Manfaat teoritis

Penelitian ini diharapkan dapat meningkatkan wawasan keilmuan mengenai analisis *cluster* menggunakan metode *ensemble*.

2. Manfaat praktis

Hasil penelitian ini diharapkan dapat memberikan informasi bagi BPRS mengenai bank mana saja harus diberi perhatian khusus.

