

# 1. Pendahuluan

## 1.1 Latar Belakang

Perkembangan dunia teknologi informasi yang terjadi saat ini secara tidak sadar telah memicu terjadinya divergensi bidang pemanfaatan teknologi informasi itu sendiri. Banyak bidang baru yang sebelumnya tidak terpikirkan untuk tersentuh oleh teknologi informasi, sekarang lambat laun mulai tersentuh akan bidang yang tidak lepas dari pengolahan data tersebut. Berangkat dari satu azas yang dinamakan optimalisasi dan peningkatan performa, masing-masing ahli dalam bidang teknologi yang berbeda mencoba untuk memformulasikan permasalahan masing-masing menjadi suatu pola keterjadian dari hasil pengukuran fenomena-fenomena yang menjadi faktor pendukungnya. Hasil pengukuran inilah yang kita kenal dengan nama data.

Dalam proses memformulasikan pola keterjadian, data tersebut diolah dengan teknik *data mining*, yaitu suatu teknik untuk menemukan suatu pola khusus pada kumpulan data untuk kemudian dianalisis lebih lanjut sehingga dapat ditarik suatu kesimpulan yang diinginkan. Salah satu metode yang dapat digunakan untuk melakukan klasifikasi data adalah metode *Classification Based on Predictive Association Rules* (CPAR). CPAR memiliki beberapa kelebihan seperti proses pembangkitan *rule* yang berbeda dan lebih efisien dibanding proses klasifikasi berbasis *rule* biasa. Namun disamping kelebihan tersebut, CPAR memiliki kelemahan yaitu bahwa CPAR tidak dapat mengatasi distribusi data yang tidak seimbang (*imbalance data set*) yang akan berakibat langsung pada pembentukan *rule* dan hasil klasifikasi [3].

Untuk mengatasi permasalahan diatas, pada tugas akhir ini akan dilakukan implementasi metode CPAR dan *Class Weighting Adjustment* terhadap studi kasus klasifikasi hari hujan. Metode CPAR ini akan dilakukan sebagai pembangkit *rule*, sedangkan metode *Class Weighting Adjustment* akan diterapkan untuk dianalisis kemampuannya dalam mengatasi kasus *imbalance data sets*. Data yang digunakan untuk klasifikasi hari hujan adalah data klimatologi yang didapat dari BMKG Stasiun Geofisika Kelas I Bandung selama lima tahun.

Klasifikasi hari hujan dipilih sebagai studi kasus karena cuaca baik jangka panjang maupun jangka pendek berperan sangat vital bagi seluruh elemen masyarakat terutama dalam hal perencanaan kegiatan. Tidak dapat dipungkiri lagi bahwa cuaca sangat mempengaruhi rencana dan aktivitas masyarakat secara langsung, baik masyarakat dengan profesi tertentu seperti petani, *event organizer*, maskapai penerbangan, maupun masyarakat pada umumnya. Selain itu klasifikasi terhadap keterjadian cuaca, khususnya hujan dilakukan karena pada penentuan cuaca dapat dilakukan dengan memperhatikan pola cuaca yang sudah terjadi dengan kondisi cuaca yang sedang terjadi [16]. Sifat yang membandingkan pola inilah menegaskan bahwa dibutuhkan suatu klasifikasi keadaan mengenai hubungan faktor fenomena alam yang terjadi terhadap keterjadian cuaca, yang dalam penelitian ini dispesifikasikan menjadi hari terjadi hujan dan hari tidak terjadi hujan. Dengan klasifikasi hari hujan ini maka diharapkan akan dapat membantu pengenalan pola terjadinya hujan.

## **1.2 Perumusan Masalah**

Dari latar belakang diatas, dapat dirumuskan beberapa *detail* permasalahan, yaitu:

1. Bagaimana mengimplementasikan metode CPAR untuk membangkitkan *rule* pada kasus klasifikasi hari hujan?
2. Bagaimana mengimplementasikan *Class Weighting Adjustment* pada *rule* yang telah dibangkitkan CPAR untuk mengetahui pengaruh *Class Weighting Adjustment* dalam menangani *imbalance data*?
3. Bagaimana performansi metode CPAR dan *Class Weighting Adjustment*?
4. Bagaimana pengaruh parameter input pada CPAR dan *Class Weighting Adjustment* terhadap akurasi?

## **1.3 Batasan Masalah**

Batasan masalah dalam penyusunan tugas akhir ini adalah sebagai berikut:

1. Hasil dari klasifikasi cuaca ini adalah variabel diskrit, yaitu akan terjadi hujan atau tidak hujan pada hari tersebut berdasarkan data yang tersedia pada hari itu juga.
2. Hasil dari klasifikasi cuaca ini tidak menyatakan curah atau intensitas hujan yang akan terjadi.
3. *Training* dan *testing sets* yang diinputkan kedalam sistem diasumsikan sudah melewati proses *preprocessing* terlebih dahulu.
4. Data yang digunakan adalah data klimatologi dari BMKG Stasiun Geofisika Kelas I selama 5 tahun (2005-2009) sebanyak 1825 *record* dengan jumlah kelas positif hujan 999 *record*, dan kelas negatif hujan sebanyak 826 *record*.
5. Data klimatologi ini mengandung beberapa variabel seperti rata-rata suhu udara ( $^{\circ}\text{C}$ ), rata-rata kelembaban nisbi (%), rata-rata kecepatan angin (Knot), arah angin terbanyak, dan kelas yang menyatakan hujan atau tidak hujan.
6. Performansi yang akan diukur adalah akurasi, *precision*, *recall* dan *f-measure*.

#### **1.4 Tujuan**

Tujuan yang ingin dicapai dalam penyusunan tugas akhir ini adalah :

5. Mengimplementasikan metode CPAR untuk membangkitkan *rule* pada kasus klasifikasi hari hujan.
6. Mengimplementasikan *Class Weighting Adjustment* pada *rule* yang telah dibangkitkan CPAR untuk mengetahui pengaruh *Class Weighting Adjustment* dalam menangani *imbalance data*.
7. Menganalisis performansi metode CPAR dan *Class Weighting Adjustment*.
8. Menganalisis pengaruh parameter input pada CPAR dan *Class Weighting Adjustment* terhadap akurasi.

#### **1.5 Metodologi penyelesaian masalah**

Metode penelitian yang dilakukan dalam menyelesaikan tugas akhir ini adalah sebagai berikut:

1. Pengumpulan data dan studi literatur  
Tahap awal dari penelitian ini adalah melakukan studi literatur dan pengumpulan data. Data yang dimaksud adalah data hasil pengamatan cuaca pada stasiun pengamatan Bandung. Sedangkan tahap studi literatur dilakukan dengan mengumpulkan informasi yang berkaitan dengan kajian yang meliputi CPAR, *Class Weighting Adjustment*, *data mining*, dan pemrograman Matlab. Informasi ini didapat dari jurnal-jurnal ilmiah, buku, *paper*, artikel dari internet, dan lain-lain.
2. Perancangan Sistem  
Tahap ini merupakan tahap untuk merancang sistem berdasarkan hasil kebutuhan. Dimana kita merancang model data, pengolahan data, alur proses, dan *interface* yang akan dibuat dalam mengimplemetasikan CPAR dan *Class Weighting Adjustment* pada klasifikasi hari hujan.
3. Implementasi Sistem  
Tahap ini adalah tahap pembuatan sistem berdasarkan pada tahap perancangan sistem dengan mengimplementasikan CPAR *Class Weighting Adjustment* pada studi kasus klasifikasi hari hujan. Bahasa pemrograman yang digunakan dalam mengimplementasikan sistem adalah Matlab.
4. Pengujian Sistem  
Tahap ini merupakan tahap untuk menguji sistem yang telah dibuat dengan berbagai macam kompleksitas yang relevan dengan sistem. Pengujian sistem ini bertujuan untuk mengetahui jika terdapat kesalahan yang mungkin terjadi saat tahap implementasi sistem dan mengevaluasi sistem agar sesuai dengan tujuan sistem tersebut dibuat.
5. Analisis Hasil  
Tahap ini adalah tahap mengevaluasi hasil keluaran sistem, yaitu berupa hasil klasifikasi, tingkat akurasi, performansi, dan lain-lain.
6. Dokumentasi  
Tahap akhir dari proses penyusunan tugas akhir ini adalah penyusunan laporan atau dokumentasi secara lengkap dan menyeluruh dari semua kegiatan pengimplementasian CPAR dan *Class Weighting Adjustment* untuk klasifikasi hari hujan.