

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Optimasi adalah proses menyelesaikan suatu masalah tertentu supaya berada pada kondisi yang paling menguntungkan dari suatu sudut pandang tertentu. Masalah yang harus diselesaikan berkaitan erat dengan data-data yang dapat dinyatakan dalam satu atau beberapa variabel (Putri Buani, 2016). Optimasi juga dapat diartikan suatu usaha atau kegiatan untuk mendapatkan hasil terbaik dengan persyaratan yang diberikan (Wanto, 2018). Proses optimasi yang dilakukan memiliki objek yang jelas artinya terdapat objek-objek yang dioptimasi. Salah satu yang dapat dijadikan objek pada penelitian optimasi adalah algoritma klasifikasi. Ada beberapa metode yang dapat digunakan untuk mengoptimasi atau meningkatkan akurasi dari suatu klasifikasi, salah satunya metode *ensemble*.

Metode *ensemble* merupakan teknik yang menggabungkan prediksi dari berbagai model untuk memberikan prediksi yang lebih akurat (Brownlee, 2017). Metode *ensemble* terdiri dari *boosting* dan *bagging*. Beberapa penelitian sebelumnya melakukan optimasi pada klasifikasi menggunakan metode *ensemble*, seperti penelitian yang mengoptimasi klasifikasi menggunakan *bagging* pada *naïve bayes* (XIAO-DONG, et al., 2010), *neural network* (Shobary, 2018) dan algoritma C4.5 (Muslim, et al., n.d.), sementara itu penelitian lainnya menggunakan *Boosting* atau *AdaBoost* pada *naïve bayes*, *SVM* dan *decision tree* (Pratama, 2016).

Penelitian prakiraan cuaca merupakan salah satu objek yang menarik untuk dikaji, karena cuaca adalah salah satu hal yang berpengaruh dalam kehidupan sehari-hari sehingga akurasi yang baik pada prakiraan cuaca menjadi sangat dibutuhkan. Prakiraan cuaca dibuat dengan mengumpulkan data kuantitatif tentang kondisi atmosfer saat ini dan menggunakan pemahaman ilmiah tentang proses atmosfer untuk memproyeksikan bagaimana atmosfer akan berkembang (Abhishek, et al., 2012).

Penelitian terkait prakiraan cuaca telah beberapa kali dilakukan dengan metode yang berbeda-beda, seperti logika fuzzy sugeno dan mamdani, *naïve bayes* dan C.45. Peneliti yang menggunakan logika fuzzy sugeno dan mamdani mendapatkan hasil akurasi terbesar 63% dan 80% (Harmoko & AZ, n.d.), sedangkan penelitian yang menggunakan metode *naïve bayes* mendapatkan akurasi sebesar 82,65% (Subhan & Fanani, 2017) dan penelitian yang menggunakan algoritma C4.5 menghasilkan akurasi sebesar 88,89% (Novandya & Oktaria, 2017). Berdasarkan deskripsi hasil penelitian sebelumnya, algoritma C4.5 dan *naïve bayes* menghasilkan nilai akurasi yang tinggi yakni 88,89% dan 82,65%.

Peramalan atau prakiraan yang bagus pada dasarnya harus memiliki nilai akurasi yang tinggi agar prediksi tersebut dapat digunakan, mendapatkan hasil yang baik dan sesuai dengan keadaan sebenarnya (Pratama, 2016). Perkembangan penelitian prakiraan cuaca selalu diharapkan memiliki akurasi yang lebih baik, cara-cara yang dapat dilakukan berupa menambah parameter input (Kurniati, et al., 2017), menggabungkan beberapa metode (Subhan & Fanani, 2017), menggunakan metode yang belum pernah digunakan sebelumnya atau menggunakan metode

optimasi (Novandya & Oktaria, 2017). Oleh, karena itu penelitian ini akan melakukan optimasi menggunakan metode *ensemble* (*bagging* dan *boosting*) pada prakiraan cuaca yang menggunakan algoritma *naïve bayes* dan *algoritma C4.5*, karena berdasarkan hasil penelitian sebelumnya algoritma *naïve bayes* dan *C4.5* memiliki nilai akurasi yang cukup tinggi yaitu 82,65% (Subhan & Fanani, 2017) dan 88,89% (Novandya & Oktaria, 2017).

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, didapatkan berbagai rumusan masalah diantaranya :

1. Bagaimana meningkatkan tingkat akurasi dengan metode *ensemble* untuk mengoptimasi objek klasifikasi prakiraan cuaca?
2. Berapa kenaikan akurasi yang dihasilkan metode *ensemble* pada klasifikasi *naïve bayes* dan *C4.5*?

## 1.3 Tujuan Penelitian

Tujuan penelitian ini adalah:

1. Melakukan optimasi menggunakan metode *ensemble* pada *Naïve Bayes* dan algoritma *C4.5*.
2. Mengukur tingkat akurasi *naïve bayes* dan algoritma *C4.5* sebelum dan setelah dioptimasi menggunakan metode *ensemble*.

#### 1.4 Batasan Masalah

Berbagai batasan masalah pada penelitian ini antara lain:

1. Dataset merupakan data klimatologi wilayah Bandung dari tahun 2009 sampai tahun 2018.
2. Parameter yang digunakan suhu, kelembaban udara, lama penyinaran matahari, kecepatan angin dan arah angin terbanyak.
3. Tools yang digunakan adalah RapidMiner.

#### 1.5 Manfaat Penelitian

Manfaat penelitian ini adalah:

1. Penelitian ini dapat dijadikan salah satu acuan untuk meningkatkan akurasi prakiraaan cuaca.
2. Mengetahui metode *ensemble* terbaik untuk mengoptimasi klasifikasi *naïve bayes* dan algoritma C4.5.

#### 1.6 Metodologi Penelitian

Terdapat beberapa tahapan penelitian yang akan dilakukan, yaitu:

1. Tahap Pendahuluan

Hal yang dilakukan pada tahap ini adalah studi literatur dan studi pustaka untuk menentukan masalah yang diangkat, batasan dan metode yang digunakan.

## 2. Tahap Pengumpulan Data

Data yang dikumpulkan merupakan data klimatologi dari BMKG Bandung.

Data yang telah dikumpulkan kemudian diproses dengan melakukan *data cleaning*, *data reduce* dan normalisasi data.

## 3. Tahap Eksperimen

Jenis eksperimen pada penelitian ini adalah *one group pretest-posttest design*. Tahap pertama yang dilakukan adalah *pretest* terlebih dahulu yakni, melakukan klasifikasi menggunakan *naïve bayes* dengan dataset dari BMKG Bandung, setelah itu dilakukan optimasi menggunakan bagging pada klasifikasi tersebut. Terakhir, dibandingkan tingkat akurasi dari klasifikasi pada *naïve bayes* sebelum dan sesudah dioptimasi.

## 4. Tahap Penyelesaian

Tahap akhir dari penelitian ini adalah pembuatan laporan yang disusun sesuai dengan ketentuan yang tercantum dalam pedoman penulisan Universitas Siliwangi Tasikmalaya, serta berisi hasil dari eksperimen yang telah dilakukan.

### 1.7 Sistematika Penelitian

#### BAB I PENDAHULUAN

Pada bagian ini dikemukakan latar belakang penelitian, rumusan masalah yang diangkat, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian dan sistematika penulisan.

## BAB II LANDASAN TEORI

Bagian ini berisikan mengenai berbagai penelitian sebelumnya (*literature review*) yang terkait optimasi dan klasifikasi, serta beberapa teori yang akan dibahas dalam penelitian ini, diantaranya, algoritma *naïve bayes*, metode *ensemble* dan RapidMiner yang merupakan tools yang digunakan dalam proses klasifikasi dan optimasi.

## BAB III METODOLOGI PENELITIAN

Bagian ini membahas jenis penelitian yang dilakukan, objek, variabel yang berpengaruh pada penelitian yang dilakukan, serta berisikan tahapan atau prosedur yang dilakukan dalam penelitian ini.

## BAB IV HASIL DAN PEMBAHASAN

Pada bagian ini dipaparkan segala hal yang dilakukan dalam penelitian sesuatu dengan metode yang digunakan. Mulai dari tahap *data preparation* sampai kepada hasil klasifikasi setelah dioptimasi, serta penarikan kesimpulan dari hasil penelitian.

## BAB V KESIMPULAN DAN SARAN

Bagian ini merupakan bagian akhir dari penelitian yang berisi mengenai kesimpulan dari penelitian yang dilakukan dan saran lain yang dapat dilakukan dalam penelitian selanjutnya.