

# 1. PENDAHULUAN

## 1.1. Latar Belakang

Kabupaten Bandung memiliki masyarakat yang mayoritas pekerjaannya adalah bercocok tanam, dengan melihat perubahan cuaca yang sangat ekstrim akhir-akhir ini, membuat hasil panen bisa terancam gagal, karena kebutuhan curah hujan merupakan salah satu hal yang dianggap penting dalam bidang pertanian untuk bercocok tanam dan juga dalam komoditas pertanian, tugas akhir ini melakukan prediksi curah hujan agar dapat membantu menentukan kapan awal bulan yang tepat untuk bercocok tanam, sehingga terhindar dari ancaman gagal panen tersebut.

Telah banyak penelitian dalam memprediksi curah hujan antara lain adalah bidang Soft Computing dan Data Mining. Pada penelitian sebelumnya dalam bidang Soft Computing, pada referensi [1,5,12,13] telah melakukan penelitian di bidang Soft Computing dengan akurasi diatas 70% untuk masing – masing algoritma yang digunakan setiap permasalahannya.

Pendekatan lain di bidang Data Mining dalam peramalan cuaca dapat menggunakan analisi klasifikasi seperti KNN (K-Nearest Neighbour), Decision Trees, Naive Bayes, dan algoritma kalsifikasi lainnya[4]. Pada referensi [3,9] telah melakukan pengujian dari algortima KNN, Decision Trees dan Naive Bayes untuk peramalan cuaca dan mendapatkan hasil akurasi 82,62 % untuk metode Decision Trees, selanjutnya pada penelitian [7] melalukan perbandingan algoritma ID3 dan C4.5 dengan informasi atribut cuaca seperti Outlook, Temperature, Humidity, Wind dengan data set sampel sebanyak 2 minggu untuk memprediksi pelaksanaan permainan golf dengan kelas “yes” atau “no”, dan menghasilkan performansi 94,15% untuk algoritma ID3 dan 96,2 % untuk algoritma C4.5.

Algoritma C4.5 adalah algoritma klasifikasi data dengan teknik pohon keputusan dan merupakan pengembangan dari algortima ID3. Algoritma C4.5 mempunyai prinsip dasar kerja yang sama dengan Algortima ID3, pengembangan tersebut adalah dapat menangani atribut kontinyu dan diskrit, dapat menangani training data dengan missing value, hasil pohon keputusan C4.5 akan dipangkas setelah dibentuk, pemilihan atribut yang dilakukan dengan menggunakan Gain Ratio [4,7,8].

Dari beberapa sumber jurnal diatas, maka pada tugas akhir ini akan dipilih teknik Data Mining dengan metode klasifikasi Decision Trees dengan algoritma C4.5 diharapkan dapat memperoleh aturan klasifikasi dari atribut tersebut dengan akurasi di atas 50%.

## 1.2. Perumusan Masalah

Permasalahan yang dapat dirumuskan berdasarkan latar belakang yang telah di uraikan diatas adalah sebagai berikut :

1. Bagaimana mengimplementasikan metode *Decision Tree* dengan algoritma C4.5 untuk peramalan data cuaca BMKG Kabupaten Bandung ?
2. Bagaimana performansi dari metode *Decision Tree* dengan algoritma C4.5 untuk peramalan cuaca ?

## 1.3. Tujuan Masalah

Tujuan dari penyelesaian Tugas Akhir ini adalah :

1. Mengimplementasikan metode Decision Trees dengan algoritma C4.5 untuk peramalan cuaca dan menggunakan data cuaca dari BMKG Kabupaten Bandung
2. Menganalisis performansi metode Decision Tree dengan algoritma C4.5 dalam memprediksi cuaca di masa yang akan datang.

## 1.4. Batasan Masalah

Batasan masalah pada Tugas Akhir ini adalah sebagai berikut :

1. Menggunakan dataset BMKG dari priode 2005 – 2016.
2. Peramalan curah hujan hanya dilakukan pada dataset BMKG untuk Kabupaten Bandung.
3. Menggunakan algoritma C4.5 dan pruning C4.5
4. Memiliki 7 atribut cuaca yaitu bulan, temperatur, curah hujan, lama penyinaran matahari, lembab nisbi, kecepatan angin, dan penguapan.

## 1.5. Metodologi Penyelesaian Masalah

Metedologi untuk menyelesaikan masalah yang telah dipaparkan adalah sebagai berikut :

1. Studi Literatur  
Mempelajari dan memahami referensi yang berhubungan dengan penelitian yang dilakukan referensi yang digunakan berasal dari jurnal, buku dan sumber lain yang ada di internet.
2. Analisis dan Perancangan Sistem  
Merancang alur sistem prediksi cuaca menggunakan *Decision Trees* dengan *Algoritma C4.5* Pengukuran performansi dari sistem dilakukan pada tahap akhir dengan metode *confusion matrix*.
3. Implementasi Sistem  
Membangun sistem yang sesuai dengan hasil rancangan sistem, setelah sistem terimplementasi, akan didapatkan hasil akhir berupa tingkat akurasi atau ukuran performansi dari sistem yang dapat dianalisis.

4. Analisis Hasil Implementasi Sistem  
Menganalisis performansi dari model *Decision Trees* dengan *Algoritma C4.5* menggunakan *confusion matrix* sebagai pengukuran performansi.
5. Pembuatan Laporan Tugas Akhir  
Membuat laporan tugas akhir berdasarkan hasil implementasi sistem dan analisis dari implementasi sistem yang dibuat.

## 1.6. Sistematika Penulisan

Untuk mempermudah pembaca dalam memahami laporan penelitian ini, penulis mengelompokkan laporan menjadi beberapa subbab dengan sistematika penulisan sebagai berikut.

1. Pendahuluan  
Bab ini menjelaskan tentang informasi umum yaitu latar belakang penelitian, perumusan masalah, batasan masalah, tujuan penelitian, metodologi penyelesaian masalah, dan sistematika penulisan.
2. Landasan Teori  
Bab ini berisikan teori yang diperlukan dalam penelitian berupa definisi dan persamaan matematika. Teori tersebut diambil dari beberapa literatur terkait di antaranya kutipan buku, *Data mining concepts and technique, conference proceeding*, dan artikel.
3. Perancangan Sistem  
Bab ini berisi tentang gambaran umum sistem dan tahapan tiap prosesnya. Selain itu, bab ini juga menjelaskan kebutuhan perangkat lunak dan keras.
4. Pengujian dan Analisis  
Dalam bab ini peneliti melakukan analisis performa yang dihasilkan dari tiap skema pengujian untuk membangun model classifier yang terbaik dalam kasus ini.
5. Kesimpulan dan Saran  
Bab ini berisi kesimpulan dan saran yang berkaitan dengan hasil analisis peneliti dan pada bab sebelumnya.