

BAB I PENDAHULUAN

1.1 Latar Belakang

Cuaca merupakan keadaan udara yang berubah setiap waktu dan daerah tertentu. Cuaca terbentuk dari gabungan dari beberapa unsur yaitu, suhu, kelembaban udara, kecepatan angin, tekanan udara, arah angin, dan lain – lainnya.

Wilayah Indonesia merupakan daerah beriklim tropis memiliki 2 musim yaitu musim hujan dan musim kemarau. Dengan iklim tropis tersebut mengakibatkan penguapan air ke udara menjadi besar, sehingga intensitas curah hujan di Indonesia menjadi tidak stabil. Curah hujan merupakan jumlah air yang jatuh ke permukaan tanah yang kemudia dihitung menggunakan satuan dalam milimeter (mm) diatas permukaan tanah yang datar.

Indonsia mengalami perubahan curah hujan dari tahun ke tahun. Beberapa provinsi di Indonsia mengalami penurunan dan peningkatan curah hujan. Salah satu provinsi di Indonesia yaitu Daerah Istimewa Yogyakarta memiliki curah hujan yang tidak stabil. Badan Pusat Statistika Yogyakarta menyatakan curah hujan pada tahun 2016 sampai 2018 menyebabkan banjir di beberapa wilayah yogyakarta. Pada tahun 2016 terjadi bencana banjir di daerah Kulon Progo, Bantul, Sleman, dan Kota Yogyakarta. Pada tahun 2017 terjadi penurunan curah hujan sehingga bencana banjir menurun. Pada tahun 2018 mengalami peningkatan curah hujan sehingga bencana banjir meningkat. Tingginya curah hujan dapat menyebabkan banjir di daerah tertentu.

Terdapat bebeberapa metode yang dapat digunakan mengklasifikasi curah hujan. Metode *Naive Bayes Classifier* dan *Support Vector Machine*. *Naive Bayes*

Classification merupakan salah satu algoritma *supervised learning* yang mudah dan efektif mengatasi derau data seperti atribut yang kurang atau tidak relevan. *Naive Bayes Classifier* masih mempunyai kekurangan dalam tingkat akurasi yang dihasilkan. Sedangkan *Support Vector Machine* (SVM) juga merupakan algoritma *supervised learning*. Kemampuan *Support Vector Machine* (SVM) menemukan *hyperplane* terbaik menjadikan algoritma SVM memiliki tingkat generalisasi yang tinggi. Kelemahan yang dimiliki SVM yaitu masih sulit digunakan untuk data yang berjumlah besar.

Pada penelitian ini dilakukan perbandingan antara dua algoritma klasifikasi *Naive Bayes Classifier* dan *Support Vector Machine* dari segi akurasi dan kecepatan proses dengan menggunakan fitur seleksi untuk mengatasi kekurangan dari *Naive Bayes Classifier* dan *Support Vector Machine*.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah, dapat diperoleh rumusan masalah sebagai berikut.

1. Bagaimana kondisi curah hujan di Yogyakarta pada Tahun 2016 – 2018?
2. Bagaimana hasil perbandingan akurasi antara metode *Naive Bayes Classifier* dan *Support Vector Machine* sebelum menggunakan *feature selection*?

3. Bagaimana hasil perbandingan akurasi antara metode *Naive Bayes Classifier* dan *Support Vector Machine* sesudah menggunakan *feature selection*?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut.

1. Data yang digunakan adalah data BMKG dari tahun 2016 sampai 2018 di Provinsi Yogyakarta.
2. Metode yang digunakan adalah metode *Naive Bayes Classifier* dan *Support Vector Machine*.
3. *Software* yang digunakan adalah *Software R*.
4. Untuk semua variabel yang digunakan merupakan data hasil rata – rata yang diperoleh dari website bmgk.go.id.

1.4 Tujuan Penelitian

Penelitian ini bertujuan sebagai berikut.

1. Mendapatkan hasil perbandingan antara metode *Naive Bayes Classifier* dan *Support Vector Machine*.
2. Mengetahui kondisi curah hujan di Yogyakarta pada Tahun 2016 – 2018.

1.5 Manfaat Penelitian

Penelitian ini dibuat dengan memiliki manfaat. Adapun manfaat penelitian ini adalah sebagai berikut.

1. Dapat mengetahui penggunaan Metode Naive Bayes Classifier dan Support Vector Machine.
2. Dapat mengetahui hasil perbedaan Metode Naive Bayes Classifier dan Support Vector Machine dengan *feature selection* dan tanpa *feature selection*.
3. Penelitian ini dapat memberikan informasi mengenai karakteristik status hujan, tidak hujan, dan hujan ekstrem yang dapat digunakan oleh BMKG Yogyakarta.

1.6 Metode Penelitian

Metode penelitian yang digunakan adalah sebagai berikut :

1.6.1 Metode pengumpulan Data

Metode pengumpulan data yang dilakukan antara lain :

1. Mengumpulkan data cuaca yang tersedia di website resmi BMKG
Data cuaca yang digunakan diambil dari data online dari web resmi BMKG <http://dataonline.bmkg.go.id/home>
2. Metode studi pustaka
Metode yang dilakukan dengan mengumpulkan data – data dari berbagai literatur seperti buku, artikel dan jurnal ilmiah yang terkait dengan topik penelitian.

1.62 Metode Testing

Metode *testing* yang digunakan untuk pengujian menggunakan dengan *software R*.

1.7 Sistematika Penulis

Penelitian ini disusun dengan sistematika penulisan sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini terdiri dari latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan yang digunakan untuk menyusun penelitian.

BAB II LANDASAN TEORI

Bab ini membahas mengenai tinjauan pustaka yang berisi dasar teori. Dan berisi tentang metode yang digunakan dalam penelitian.

BAB III METODELOGI PENELITIAN

Bab ini berisi tentang alat, bahan serta alur penelitian yang dilakukan. Memisahkan data *training* dan data *testing* yang akan diuji.

BAB IV ANALISIS DAN PEMBAHASAN

Bab ini berisi pemaparan hasil dari pengujian data yang digunakan.

BAB V PENUTUP

Bab ini berisi kesimpulan dari penelitian serta saran untuk memperbaiki kekurangan pada penelitian.