

BAB I PENDAHULUAN

1.1 Latar Belakang

Perubahan kondisi cuaca yang berlangsung cepat dan tidak menentu menjadi tantangan dalam berbagai bidang kehidupan. Berbagai bidang seperti pertanian, transportasi, pariwisata, dan pengelolaan bencana alam sangat membutuhkan informasi cuaca yang akurat untuk mendukung pengambilan keputusan dalam aktivitas sehari-hari. Mulai dari keputusan jadwal penanaman, pengiriman logistik, perencanaan perjalanan, dan penanganan dini terhadap potensi bencana alam yang bergantung pada kondisi atmosfer yang sedang terjadi. Sistem cuaca sendiri merupakan fenomena atmosfer yang bersifat kompleks dan dinamis, karena melibatkan berbagai parameter meteorologis seperti suhu udara, kelembapan, tekanan, kecepatan angin, hingga radiasi matahari. Hubungan antara parameter tersebut tidak linear dan sulit untuk diprediksi. Meskipun saat ini sudah tersedia data cuaca dalam jumlah besar melalui stasiun cuaca, satelit, dan radar cuaca. Dalam era *big data* ini, pendekatan berbasis *machine learning* banyak diterapkan dalam memprediksi cuaca dengan kemampuannya dalam mengenali pola-pola dalam data yang bersifat kompleks dan tidak linear. Hal ini menjadi solusi untuk meningkatkan akurasi prediksi cuaca, terutama tantangan dalam perubahan iklim dan ketidakpastian atmosfer yang semakin meningkat [1].

Dalam pemodelan prediksi cuaca tidak hanya terletak pada kompleksitas sistem atmosfer, tetapi juga pada kualitas data meteorologi yang tersedia. Data yang diperoleh seringkali mengandung *missing values*, inkonsistensi, maupun gangguan pencatatan yang dapat menurunkan performa model prediksi [2]. Upaya untuk meningkatkan kualitas data, misalnya melalui teknik *imputasi* dan *praproses* data, telah terbukti berpengaruh signifikan terhadap hasil peramalan cuaca yang lebih akurat [3]. Dengan demikian, tantangan utama terletak pada bagaimana membangun model yang tidak hanya mampu mengenali pola yang kompleks, tetapi juga tangguh terhadap keterbatasan data yang ada.

Untuk mengatasi berbagai tantangan dalam memprediksi cuaca, penggunaan *machine learning* telah menunjukkan potensi yang besar dalam meningkatkan akurasi prediksi cuaca. Berbagai algoritma *single classifier* seperti Decision Tree dan Naive Bayes telah diterapkan dalam penelitian prediksi cuaca dengan hasil yang cukup baik [4]. Support Vector Machine (SVM) juga menunjukan performa yang baik dalam penelitian

yang serupa [5]. Masing masing memiliki kelebihan dan kelemahan dalam menangani data cuaca yang kompleks. Menurut [6] Decision Tree mudah dipahami namun cenderung *overfitting*, sementara itu Naive Bayes sederhana dan cepat namun mengasumsikan bahwa setiap fitur bersifat independen, yang tidak selalu sesuai dengan kondisi cuaca yang saling berkaitan. Dan seperti [7] SVM efektif untuk data yang berdimensi tinggi namun sensitif pada parameter. Keterbatasan dari masing masing algoritma ini mendorong pengembangan *ensemble learning*[8].

Penelitian ini melakukan pendekatan untuk meningkatkan akurasi prediksi cuaca menggunakan *ensemble learning*, yaitu metode yang menggabungkan beberapa model untuk menghasilkan prediksi yang lebih akurat dan stabil. Salah satu metode *ensemble learning* yang digunakan dalam penelitian ini adalah Stacking Classifier, yang dipilih karena kemampuannya menggabungkan kekuatan dari berbagai algoritma *machine learning* yang berbeda melalui *meta-model* untuk menghasilkan prediksi akhir yang lebih optimal. Fokus penelitian berada pada tugas klasifikasi, karena kategori cuaca seperti hujan atau tidak hujan merupakan informasi paling mendasar dan penting dalam kehidupan sehari-hari. Pendekatan stacking dianggap cocok untuk menghadapi tantangan prediksi cuaca yang kompleks, karena mampu menangani data meteorologi yang bersifat *non-linear*, memiliki banyak variabel, serta rentan terhadap *missing values* dan *noise* [9].

Oleh karena itu, dengan melakukan analisis terhadap metode *Stacking Classifier*, penelitian ini bertujuan untuk membangun model prediksi yang lebih akurat, stabil, dan dapat diandalkan. Kontribusi dari penelitian ini tidak hanya diharapkan memberikan solusi prediksi cuaca, tapi juga diharapkan dapat menjadi acuan atau referensi dalam pengembangan sistem prediksi cuaca kedepannya. Serta mendukung berbagai sektor-sektor penting yang bergantung pada informasi cuaca yang akurat.

1.2 Rumusan Masalah

1. Apakah metode ensemble learning lebih baik dibandingkan dengan metode *single classifier* dalam memprediksi cuaca?
2. Seberapa pengaruh metode *Stacking Classifier* dalam mengenali pola cuaca berdasarkan variabilitas parameter meteorologi?
3. Sejauh mana proses hyperparameter tuning mempengaruhi performa model *Stacking Classifier* dalam memprediksi cuaca?

1.3 Batasan Masalah

1. Prediksi yang dilakukan pada penelitian ini dibatasi pada prediksi cuaca berbasis klasifikasi, seperti kategori hujan atau tidak, atau kategori cuaca lainnya tergantung pada data yang tersedia. Penelitian ini tidak membahas prediksi dalam bentuk numerik (regresi), seperti suhu dalam berapa derajat atau kecepatan angin dalam kecepatan tertentu.
2. Model yang digunakan hanya metode stacking, serta beberapa model single classifier sebagai pembandingan. Algoritma lain diluar metode tersebut tidak akan dibahas secara mendalam.
3. Data yang digunakan dalam penelitian ini diperoleh dari sumber terbuka seperti Kaggle. Metode ini tidak melibatkan pengambilan data langsung dari alat ukur cuaca seperti sensor atau stasiun cuaca.
4. Faktor-faktor seperti kondisi seperti permukaan tanah, perbedaan ketinggian antar wilayah, dan variabel luar lainnya yang sulit diukur atau dimasukkan kedalam model, diasumsikan tetap atau tidak dihitung secara langsung di penelitian ini. Hal ini berfokuskan agar penelitian tetap pada analisis data cuaca yang relevan.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini adalah untuk mengevaluasi kinerja ensemble learning dalam melakukan prediksi cuaca berbasis klasifikasi. Selain itu, dilakukan perbandingan hasil prediksi dengan beberapa metode *single classifier*, untuk mengetahui apakah pendekatan kombinasi model mampu memberikan performa yang lebih baik. Fokus lainnya adalah mengidentifikasi metode ensemble mana yang paling optimal, dalam menangani data yang meteorologi yang kompleks. Analisis juga dilakukan untuk melihat sejauh mana metode tersebut mengenali pola cuaca berdasarkan variabilitas parameter meteorologi.

1.5 Manfaat Penelitian

Manfaat yang dapat diperoleh dari penelitian ini sebagai berikut:

1. Memberikan kontribusi terhadap pengembangan ilmu di bidang data science dan meteorologi, khususnya dalam pemanfaatan metode ensemble learning untuk pemodelan prediksi cuaca. Hasil analisis ini diharapkan dapat menjadi referensi ilmiah bagi penelitian lain yang sedang mengembangkan model prediksi cuaca berbasis machine learning untuk data cuaca.

2. Penelitian ini bisa dapat digunakan oleh instansi atau lembaga yang bergerak di bidang informasi cuaca untuk meningkatkan hasil akurasi prediksi cuaca. Model yang dihasilkan dapat membantu dalam pengambilan keputusan yang tepat, baik dalam sektor pertanian, transportasi, pariwisata, dan pengelolaan bencana alam.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN, berisi latar belakang yang memberikan gambaran awal mengenai pentingnya prediksi cuaca dan peran metode *Stacking Classifier*. Mencakup rumusan masalah yang terjadi, manfaat penelitian secara teoritis dan praktis, dan arah yang ingin dicapai dalam penelitian.

BAB II TINJAUAN PUSTAKA, berisi uraian dasar teori dan literatur yang relevan dengan penelitian. Pembahasan mencakup pengertian dan karakteristik data cuaca, konsep dasar machine learning dan klasifikasi, serta penjelasan tentang metode prediksi cuaca single classifier maupun metode ensemble secara teori, termasuk teknik-teknik seperti Bagging, Boosting, dan Stacking. Serta tinjauan terhadap penelitian-penelitian terdahulu yang relevan juga disertakan, untuk menunjukan gap yang ingin diisi dalam penelitian ini. Bab ini juga menjelaskan hubungan antara variabel yang diteliti.

BAB III METODE PENELITIAN, memaparkan langkah langkah dan pendekatan yang digunakan dalam melakukan penelitian. Penjelasan meliputi jenis penelitian, sumber dan karakteristik data yang digunakan, proses pra proses data seperti missing values dan normalisasi, pemilihan fitur, serta pembagian data pelatihan dan pengujian. Dijelaskan juga algoritma yang digunakan baik single classifier maupun metode ensemble. Serta metrik evaluasi yang digunakan seperti akurasi, precision, recall, dan F1-score. Diagram alur penelitian dan tools pemrograman juga dipaparkan dalam penelitian ini.

BAB IV HASIL DAN PEMBAHASAN, menyajikan hasil dari eksperimen yang telah dilakukan, dari hasil training dan testing model, hingga perbandingan performa antara model. Dibahas juga analisis terhadap kelebihan dan kekurangan pada masing-masing pendekatan berdasarkan metrik evaluasi. Penjelasan mencakup efektivitas ensemble

learning dibandingkan single classifier dalam mengklasifikasi cuaca, dan visual hasil seperti grafik, dan tabel performa model juga ditampilkan untuk mendukung analisis.

BAB V PENUTUP, memberi kesimpulan dari hasil utama dari penelitian, berdasarkan rumusan masalah yang sudah dibuat, disampaikan ringkasan temuan bahwa ensemble learning mampu memberikan kinerja seperti apa pada prediksi cuaca dibandingkan model individual. Bab ini juga memberikan saran untuk pengembangan penelitian selanjutnya yang diharapkan dapat berkontribusi bagi yang sedang mengembangkan ilmu meteorologi.

