

**PREDIKSI CUACA DI KABUPATEN SAMBAS
MENGUNAKAN ALGORITMA
RANDOM FOREST**

TUGAS AKHIR



OLEH:

DEDY KURNIAWAN
NPM. 191220057

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS MUHAMMADIYAH PONTIANAK
2025**

BAB I

PENDAHULUAN

1.1 Latar Belakang

Prediksi cuaca merupakan aspek penting dalam berbagai sektor, termasuk pertanian, transportasi, dan manajemen bencana. Di Kabupaten Sambas, yang terletak di Kalimantan Barat, Indonesia, perubahan iklim dan pola cuaca yang tidak menentu menuntut pengembangan model prediksi yang lebih akurat untuk membantu masyarakat dan pemerintah dalam perencanaan dan mitigasi risiko.

Cuaca yang tidak terduga dapat berdampak signifikan pada kegiatan ekonomi dan sosial. Misalnya, petani membutuhkan informasi akurat tentang curah hujan untuk menentukan waktu tanam dan panen. Selain itu, sektor pariwisata juga bergantung pada kondisi cuaca yang baik untuk menarik pengunjung. Oleh karena itu, model prediksi cuaca yang efektif sangat diperlukan untuk memberikan informasi yang tepat waktu dan akurat.

Algoritma Random Forest adalah salah satu teknik machine learning yang telah terbukti efektif dalam berbagai aplikasi prediksi, termasuk prediksi cuaca. Metode ini bekerja dengan membangun banyak pohon keputusan dan menggabungkan hasilnya untuk meningkatkan akurasi prediksi. Keunggulan Random Forest terletak pada kemampuannya untuk menangani data dengan banyak atribut dan mengurangi risiko overfitting dibandingkan dengan model lain.

Dalam konteks penelitian di Kabupaten Sambas, penggunaan algoritma Random Forest dapat dilakukan dengan langkah-langkah berikut:

1. Pengumpulan Data: Mengumpulkan data historis cuaca dari stasiun meteorologi setempat, termasuk suhu, kelembaban, kecepatan angin, dan curah hujan.
2. Pengumpulan Data: Mengumpulkan data historis cuaca dari stasiun meteorologi setempat, termasuk suhu, kelembaban, kecepatan angin, dan curah hujan.
3. Pengumpulan Data: Mengumpulkan data historis cuaca dari stasiun meteorologi setempat, termasuk suhu, kelembaban, kecepatan angin, dan curah hujan.

4. Pengumpulan Data: Mengumpulkan data historis cuaca dari stasiun meteorologi setempat, termasuk suhu, kelembaban, kecepatan angin, dan curah hujan.
5. Pengumpulan Data: Mengumpulkan data historis cuaca dari stasiun meteorologi setempat, termasuk suhu, kelembaban, kecepatan angin, dan curah hujan.

Penelitian sebelumnya menunjukkan bahwa algoritma Random Forest dapat menghasilkan akurasi yang baik dalam memprediksi curah hujan. Sebagai contoh, penelitian di Bandar Lampung menggunakan algoritma ini menunjukkan nilai MAE sekitar 0.111 hingga 0.1295 dan RMSE antara 0.2218 hingga 0.2523[1]. Hasil ini menunjukkan bahwa meskipun Random Forest mungkin tidak selalu seakurat metode lain seperti C4.5 dalam beberapa kasus, ia tetap merupakan pilihan yang solid untuk aplikasi prediksi cuaca.

Dengan menggunakan algoritma Random Forest dalam prediksi cuaca di Kabupaten Sambas, diharapkan dapat meningkatkan akurasi dan keandalan informasi cuaca yang diterima oleh masyarakat. Ini akan membantu dalam perencanaan kegiatan sehari-hari serta mitigasi risiko bencana terkait cuaca ekstrem. Pengembangan lebih lanjut dalam penelitian ini dapat berkontribusi pada sistem informasi cuaca yang lebih efektif di wilayah tersebut.

1.2 Rumusan Masalah

1. Bagaimana kinerja algoritma Random Forest dalam memprediksi kondisi cuaca di Kabupaten Sambas?
2. Faktor-faktor meteorologi apa saja yang paling berpengaruh terhadap keakuratan prediksi cuaca di wilayah tersebut?
3. Seberapa akurat model prediksi cuaca berbasis algoritma Random Forest?

1.3 Batasan Masalah

1. Penelitian ini berfokus pada implementasi data mining menggunakan Algoritma Random Forest untuk prediksi cuaca.
2. Penelitian ini difokuskan pada prediksi cuaca untuk wilayah Kabupaten Sambas saja.

3. Penelitian ini akan menggunakan dataset yang tersedia di website BMKG dengan 11 atribut.
4. Data yang digunakan sebanyak 366 data.
5. Preprocessing data dilakukan dengan menggunakan bahasa pemrograman python dan visual studio code sebagai text editor.

1.4 Tujuan Penelitian

1. Menerapkan algoritma Random Forest untuk memprediksi kondisi cuaca di Kabupaten Sambas.
2. tujuan penelitian ini juga membangun aplikasi praktis dari model prediktif yang dikembangkan, seperti penggunaan perkiraan cuaca untuk keperluan sehari-hari.

1.5 Manfaat Penelitian

1. Bagi Masyarakat dan Pemerintah Daerah: Penelitian ini dapat memberikan informasi prediksi cuaca yang lebih akurat, sehingga masyarakat dan pemerintah daerah dapat mengambil langkah preventif dalam menghadapi kondisi cuaca ekstrem yang dapat berdampak pada sektor pertanian, perikanan, dan aktivitas ekonomi lainnya.
2. Bagi Peneliti dan Akademisi: Hasil penelitian ini dapat menjadi referensi dan bahan kajian lebih lanjut dalam pengembangan metode prediksi cuaca menggunakan teknik pembelajaran mesin, khususnya algoritma Random Forest.
3. Bagi Pengembangan Teknologi Prediksi Cuaca: Model prediksi cuaca berbasis algoritma Random Forest yang dihasilkan dapat menjadi acuan untuk meningkatkan sistem prediksi cuaca yang ada, sehingga lebih adaptif dan akurat dalam menghadapi perubahan iklim yang dinamis.

1.6 Metodologi Penelitian

Penelitian dilakukan untuk memprediksi perkiraan cuaca menggunakan data mining dengan metode Random Forest. Langkah-langkah yang dilakukan

dalam penelitian ini diantaranya Identifikasi Masalah, Studi Literatur, Pengumpulan Data, Implementasi, Pengujian Sistem, dan Pengambilan Kesimpulan.

1.7 Sistematika Penulisan

Struktur penulisan penelitian ini terdiri dari beberapa bab yaitu pendahuluan, landasan teori, metode penelitian, analisis dan perancangan sistem, hasil dan pengujian, serta kesimpulan dan saran.

BAB I Pendahuluan

Pada bab pendahuluan ini menjelaskan latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penelitian.

BAB II Landasan Teori

Pada bab ini berisikan tinjauan pustaka mengulas teori-teori terkait Collaborative Filtering serta mengkaji penelitian-penelitian terdahulu

BAB III Metode Penelitian

Pada bab ini membahas rancangan penelitian, pengumpulan dan analisis data, serta langkah-langkah pengembangan sistem prediksi.

BAB IV Analisis dan Perancangan Sistem

Pada bab ini membahas analisis kebutuhan sistem berdasarkan hasil pengumpulan dan analisis data. Sistem rekomendasi ini dikembangkan dalam bentuk platform website.

BAB V Hasil dan Pengujian

Pada bab ini dilakukan untuk menguji kinerja dan fungsionalitas sistem rekomendasi. Selanjutnya di uji kembali apakah metode berhasil sesuai yang diharapkan untuk menarik kesimpulan dan saran.

BAB VI Kesimpulan dan Saran

Pada bab ini merumuskan kesimpulan dari hasil penelitian dan memberikan saran-saran untuk pengembangan lebih lanjut.

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Tinjauan pustaka dilakukan untuk mengarahkan penelitian ini dengan memastikan fokus yang tepat dan menghindari duplikasi dari penelitian sebelumnya. Tinjauan pustaka juga mencakup ulasan tentang penelitian terdahulu yang relevan dengan penelitian yang akan dilakukan, yang akan menjadi acuan dan referensi bagi penulis. Terdapat beberapa penelitian yang membahas mengenai cuaca dan beberapa penelitian yang membahas algoritma Random Forest.

1. Penelitian oleh Zian Asti Dwiyantri, Cahyo Prianto pada tahun 2023 dengan judul “Prediksi Cuaca Kota Jakarta menggunakan Metode Random Forest: Studi Optimalitas”. Tujuan dari penelitian ini adalah untuk membuat prediksi cuaca di Kota Jakarta dengan menggunakan metode Random Forest dan data cuaca historis yang dapat diandalkan dari website OpenData Jakarta. Hasil evaluasi menunjukkan bahwa model Random Forest memberikan prediksi cuaca yang baik dengan akurasi, presisi, dan recall sebesar 0,71, skor F1 sebesar 0,70, dan ROC-AUC sebesar 0,92. [2].
2. Penelitian oleh Ghaitsa Amany Mursianto, Isma’il Muhammad Falih, Muhammad Irfan, Tiara Sakinah, Desta Sandya Prasvita pada tahun 2021 dengan judul “Perbandingan Metode Klasifikasi Random Forest dan XGBoost Serta Implementasi Teknik SMOTE pada Kasus Prediksi Hujan”. Tujuan penelitian ini adalah untuk mengklasifikasikan prediksi hujan pada hari-hari berikutnya dengan menggunakan metode klasifikasi Random Forest. Sistem prediksi cuaca yang telah kami buat mendapatkan klasifikasi Random Forest dengan resampling dengan akurasi 95,59%, tetapi XGBoost mendapatkan akurasi tertinggi pada klasifikasi tanpa resampling dengan akurasi 94,34% [3].

3. Penelitian oleh Agung Hot Iman, Fransisco Ready Permana, Gito Putro Wardana, Raihan Kemmy Rachmansyah, Mayanda Mega Santoni pada tahun 2022 dengan judul “Perbandingan Algoritma Klasifikasi Random Forest dan Extreme Gradient Boosting pada Dataset Cuaca Provinsi DKI Jakarta Tahun 2018”. Setelah melewati tahap preprocessing, analisis data ekstensif, pengecekan label, dan pelatihan data dilakukan untuk mendapatkan hasil yang akurat. Hasil penelitian menunjukkan bahwa, jika digunakan pada data cuaca Provinsi DKI Jakarta tahun 2018, algoritma Random Forest akan lebih baik daripada algoritma Extreme Gradient Boosting. Ini karena waktu pemrosesan yang lebih cepat dan akurasi yang dihasilkan cukup tinggi yaitu 68% [4].
4. Penelitian oleh Risanti, Widyaningrum Indrasari, Haris Suhendar pada tahun 2024 dengan judul “Analisis Model Prediksi Cuaca Menggunakan Support Vector Machine, Gradient Boosting, Random Forest, Dan Decision Tree”. Penelitian ini menggunakan model support vector machine, gradient boosting, random forest, dan decision tree. Perbandingan antara data pelatihan dan tes adalah 70:30. Hasil penelitian menunjukkan bahwa akurasi model support vector machine, gradient boosting, random forest, dan decision tree masing-masing sebesar 0.1697, 0.6696, 0,7918, 0,8416, dan 0,8280. Ini menunjukkan bahwa random forest memiliki pengaruh paling besar terhadap hasil prediksi cuaca dengan hasil akurasi 0.8416 [5].
5. Penelitian oleh M. Imam Whidyarto, Casi Setianingsih, Wendi Harjupa, Risyanto pada tahun 2023 dengan judul “Prediksi Curah Hujan Dari Data Satelit Himawari-8 Menggunakan Metode Random Forest”. Metode Random Forest (RF) menggunakan data netCDF untuk memprediksi ketinggian awan dan hujan. Ini menggunakan titik acuan, yaitu kilometer (km). Akurasi tertinggi adalah 100% dalam pengujian data 120 data. [6].

2.2 Data Mining

Data mining adalah serangkaian prosedur yang digunakan untuk mengidentifikasi pola dan informasi yang berguna dari suatu basis data yang sebelumnya tidak diketahui secara manual. Data mining muncul pada tahun 1990-an dan digunakan secara efektif dan akurat untuk menemukan hubungan antar data dan menemukan pola dan informasi yang berguna untuk pengelompokan data dalam satu atau lebih cluster sehingga objek-objek dalam cluster memiliki tingkat kemiripan yang tinggi satu sama lain. Bagian dari proses penemuan pengetahuan dari basis data, Knowledge Discovery in Databases, adalah data mining. [7].

Menurut Kusriani dan Taufiq Luthfi (2009), pemrosesan data dipilah menjadi beberapa kategori berdasarkan fungsi yang dapat dilakukan, seperti:

1. Deskripsi

Kadang-kadang, analisis dan penelitian hanya ingin mencari cara untuk menunjukkan pola dan kecenderungan dalam data.

2. Estimasi

Kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori, estimasi hampir sama dengan klasifikasi. Model dibangun dengan record lengkap yang menyediakan nilai dan variabel target sebagai nilai prediksi; kemudian, pada peninjauan berikutnya, nilai estimasi variabel target didasarkan pada nilai variabel prediksi.

3. Prediksi

Prediksi dan klasifikasi hampir sama, kecuali prediksi nilai dari hasil di masa mendatang.

4. Klasifikasi

Dalam klasifikasi, variabel target kategori digunakan. Sebagai contoh, penggolongan pendapat dapat dikategorikan menjadi tiga kategori: pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

5. Pengklusteran

Kumpulan catatannya yang mirip satu sama lain dan tidak sama dengan catatannya dalam kluster lain disebut "klusteran". Kluster adalah kumpulan catatannya yang dikelompokkan, diamati, atau diperhatikan untuk membentuk kelas catatannya yang memiliki kemiripan.

6. Asosiasi

Menemukan fitur yang muncul dalam satu waktu adalah tugas bersama dalam datamining. Analisis keranjang belanja lebih dikenal di dunia bisnis [8].

2.3 Random Forest

Seperti namanya, Random forest terdiri dari banyak pohon keputusan individu yang bekerja sama sebagai kelompok. Setiap pohon random forest individu menghasilkan prediksi kelas, dan kelas dengan suara terbanyak dimasukkan ke dalam prediksi model. Visualisasi model Random forest membuat perkiraan. Konsep dasar Random forest sederhana tetapi kuat. [9]. Algoritma pohon random (RF) menggunakan pemisah biner rekursif untuk mencapai node akhir dalam struktur pohon keputusan, yang pada dasarnya terdiri dari pohon regresi dan klasifikasi. Hutan acak menghasilkan banyak pohon independen dengan menggunakan teknik bootstrap dan subset acak dari data pelatihan dan variabel input. Agregasi bootstrap, juga disebut bagging, melibatkan penggunaan teknik bootstrap atau pengambilan sampel berkali-kali untuk membuat set instruksi yang berbeda dengan menggantikan data instruksi asli. Setiap data pelatihan yang dikumpulkan selama proses ini digunakan untuk membuat pohon klasifikasi. Hasil klasifikasi adalah agregasi suara terbanyak, dan hasil regresi adalah rata-rata. Berikut adalah struktur umum dari random forest yang ditunjukkan Gambar 1 [10].

Langkah-langkah Algoritma Random Forest:

2.3.1 Persiapan Data

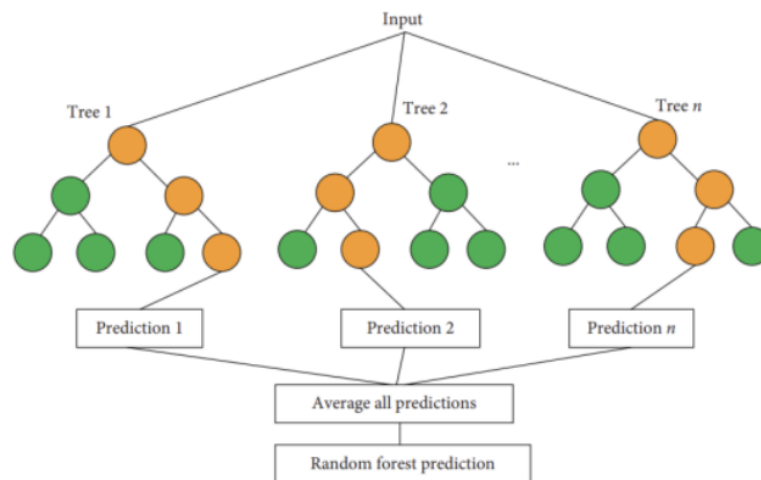
- Kumpulkan dataset yang berisi fitur (variabel independen) dan label (variabel target). Misalnya, untuk prediksi cuaca, fitur dapat mencakup suhu, kelembaban, curah hujan, dll.
- Bagi dataset menjadi data latih (training set) dan data uji (testing set).

2.3.2 Membangun Pohon Keputusan (Decision Trees)

- tentukan jumlah pohon keputusan (misalnya, 100 atau 500 pohon) yang akan digunakan dalam model Random Forest.
- Untuk setiap pohon keputusan:
 - Lakukan bootstrap sampling, yaitu memilih sampel secara acak dari data latih dengan penggantian. Dengan demikian, setiap pohon dibangun menggunakan subset data yang berbeda.
 - Pilih secara acak subset fitur dari semua fitur yang tersedia. Ini dilakukan untuk setiap pembagian simpul pada pohon untuk mengurangi korelasi antara pohon.
 - Bangun pohon keputusan dengan menggunakan subset data dan fitur yang dipilih. Setiap pohon tumbuh secara penuh tanpa memangkas cabangnya.

2.3.3 Melakukan Prediksi

- Untuk membuat prediksi, setiap pohon memberikan prediksi sendiri untuk data uji atau data baru.
- Prediksi akhir untuk klasifikasi ditentukan dengan metode voting mayoritas (prediksi yang paling sering muncul di antara semua pohon).
- Untuk regresi, prediksi akhir ditentukan dengan menghitung rata-rata prediksi dari semua pohon.



Gambar 1 Struktur umum model *Random Forest* [10]

2.4 Visual Studio Code

Visual Studio Code adalah editor teks Microsoft yang sederhana dan kuat yang kompatibel dengan berbagai sistem operasi, termasuk Linux, Mac, dan Windows. Ini mendukung bahasa pemrograman seperti JavaScript, Typescript, dan Node.js secara langsung, dan mendukung bahasa pemrograman lainnya dengan plugin yang dapat diunduh dari marketplace Visual Studio Code, seperti C++, C#, Python, Go, Java, PHP, dan lainnya [11].

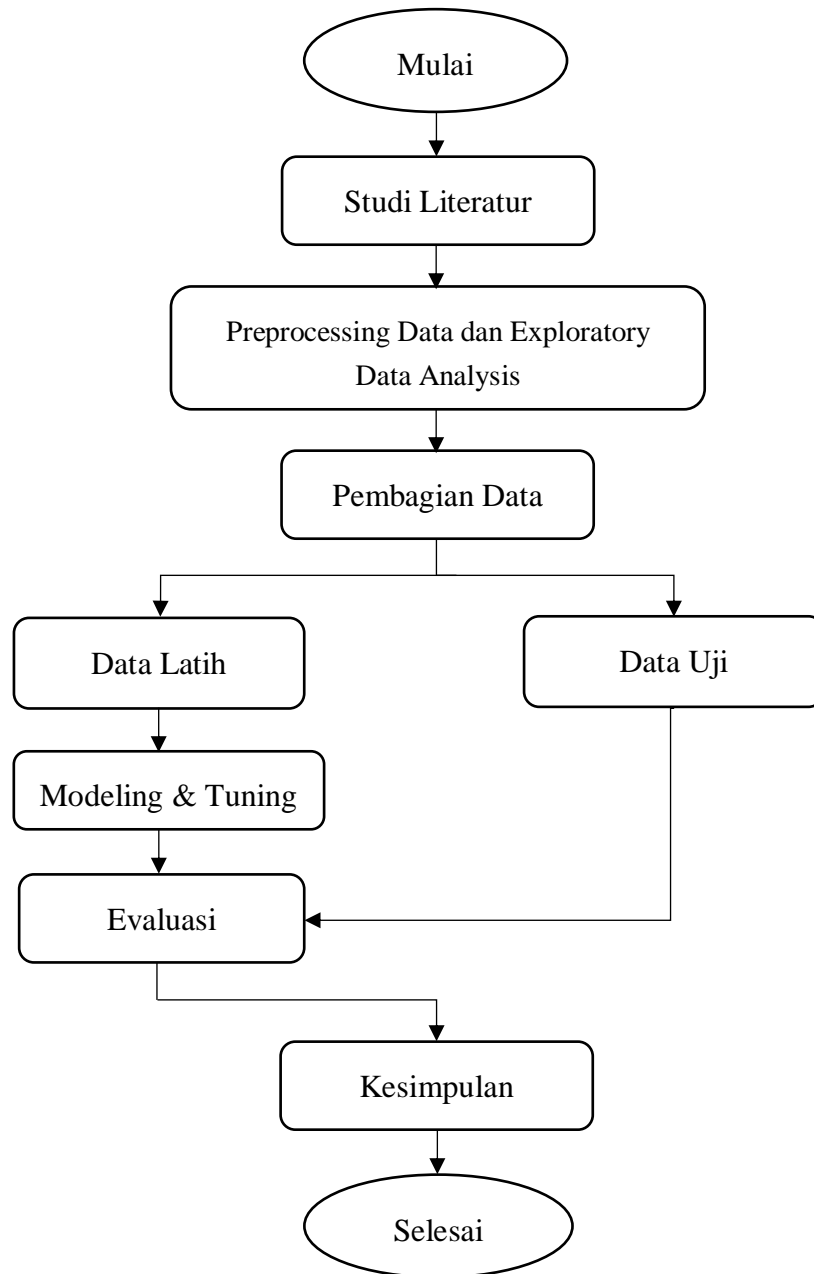
2.5 Python

Python adalah bahasa pemrograman bersifat open source yang dioptimalkan untuk kualitas software, produktivitas developer, portabilitas program, dan integrasi komponen. Python saat ini merupakan salah satu dari empat atau lima bahasa pemrograman yang paling populer di dunia dan digunakan untuk membangun berbagai jenis perangkat lunak, seperti skrip internet, pemrograman sistem, antarmuka pengguna, penyesuaian produk, dan penghitungan numerik programming.

Python, sebagai bahasa pemrograman, menawarkan serangkaian fitur yang dapat digunakan oleh para pengembang perangkat lunak. Berikut beberapa dari fitur-fitur yang dimiliki oleh bahasa pemrograman Python::

1. Multi Paradigm Design.
2. Open Source.
3. Simplicity.
4. Library Support.
5. Portability.
6. Extendable.
7. Scalability [12].

BAB III METODE PENELITIAN



Gambar 2 Tahapan Metode Penelitian

3.1 Studi Literatur

Pada tahapan ini peneliti melakukan studi literatur untuk mendapatkan teori-teori yang valid terkait dengan penelitian yang dilakukan sebelumnya. Sumber literatur

yang digunakan untuk dijadikan referensi antara lain jurnal, artikel, dan situs internet.

3.2 Input Dataset

Data yang digunakan untuk penelitian ini diambil dari open dataset BMKG

3.3 Preprocessing Data dan Exploratory Data Analysis

Tahapan selanjutnya, peneliti melakukan pra-proses dan exploratory data analysis pada data yang telah peroleh. Tahapan ini bertujuan untuk membersihkan data dari derau (noise), mendapatkan lebih banyak insight dari data dengan exploratory data analysis, dan mengekstrak informasi yang ada pada data dengan melakukan feature engineering.

3.4 Pembagian Data

Tahapan setelah preprocessing dan exploratory data analysis adalah pembagian data menjadi dua yaitu data latih dan data uji dengan perbandingan 80% data latih dan 20% data uji. Selanjutnya akan dipisahkan antara kolom label dan kolom fitur. Pada kolom label dilakukan pembagian secara stratifikasi atau pembagian data secara seimbang karena label yang digunakan bertipe kategorikal.

3.5 Preprocessing, Modeling, dan Tuning

3.5.1 Preprocessing

Sebelum tahapan selanjutnya, peneliti mempersiapkan data terlebih dahulu agar lebih siap untuk di proses. Peneliti melakukan tahapan pra-proses kembali setelah tahap pembagian data. Hal ini peneliti lakukan untuk menghindari kebocoran data sehingga akan mengurangi tingkat overfitting ataupun underfitting. Pada tahapan pra-proses kali ini, peneliti menggunakan Column Transformer dengan membagi kolom menjadi kolom numerik dan kolom kategorik. Untuk kolom numerik dilakukan pra-proses berupa Polynomial Feature sedangkan kolom kategorik dilakukan pra-proses berupa One-Hot Encoder dengan tujuan untuk membangun model yang sederhana yang dapat digunakan untuk pemahaman dan melakukan prediksi. Polynomial Feature adalah sebuah teknik untuk mengatasi data yang tidak linier dengan menambahkan kekuatan pada setiap fitur sebagai fitur yang baru. Polyomial Feature akan melatih model linier pada kumpulan fitur yang telah

diperluas. One-Hot Encoder sendiri adalah salah satu metode encoding yang merepresentasikan data bertipe kategori sebagai vektor biner dengan nilai integer, 0 dan 1, dimana semua elemen akan bernilai 0 kecuali satu elemen yang bernilai 1, yaitu elemen yang memiliki nilai kategori tersebut.

3.5.2 Modeling

Setelah membuat Column Transformer, Kemudian peneliti membuat Pipeline untuk menggabungkan Column Transformer yang telah dibuat sebelumnya dengan algoritma yang akan peneliti gunakan yaitu Random Forest dan Extreme Gradient Boosting. Random Forest adalah pengembangan dari metode pohon keputusan yang menggunakan beberapa pohon keputusan, dimana setiap pohon keputusan telah dilakukan pelatihan menggunakan sampel individu dan setiap atribut dipecah pada pohon yang dipilih antara atribut subset yang bersifat acak. Sementara itu, Extreme Gradient Boosting adalah metode boosting yang menggabungkan beberapa pohon keputusan dan dibangun dengan cara membuat pohon keputusan yang baru berdasarkan residual pohon keputusan sebelumnya dengan tujuan untuk menghasilkan kesalahan prediksi (error) yang lebih kecil dari model sebelumnya. Pipeline sendiri adalah sebuah metode yang dapat mengatur aliran data input dan output dari sebuah model atau kumpulan beberapa model machine learning. Salah satu keuntungan menggunakan pipeline adalah pipeline dapat memproses data mulai dari input data mentah, fitur, output, model machine learning dan parameter model, dan output prediksi.

3.5.3 Tuning

Pembuatan Column Transformer dan pipeline bertujuan untuk mempermudah proses dan memprediksi data baru dengan model telah dibuat serta memungkinkan kita dalam melakukan tuning hyperparameter berdasarkan algoritma yang ada di dalam pipeline. Tuning hyperparameter sendiri merupakan proses untuk menemukan nilai hyperparameter terbaik yang sesuai dengan data yang dimiliki sehingga dapat menghasilkan model dengan performa yang terbaik.

3.6 Evaluasi

Setelah semua tahapan selesai, selanjutnya akan dilakukan tahapan evaluasi performa dari model Random Forest dan Extreme Gradient Boosting menggunakan data uji. Evaluasi model bertujuan untuk membuat estimasi generalisasi error dari model yang dipilih, dan menguji seberapa baik kinerja model tersebut pada data baru yaitu data uji. Peneliti menggunakan confusion matrix untuk mengetahui efisiensi kinerja model agar dapat diketahui seberapa banyak data yang salah diklasifikasi pada data uji dan classification report untuk mengukur tingkat akurasi, presisi, dan recall.

DAFTAR PUSTAKA

- [1] Sudarta, *PRAKIRAAN CUACA KOTA BANDAR LAMPUNG DENGAN MENGGUNAKAN RANDOM FOREST DAN C4.5*, vol. 16, no. 1. 2022.
- [2] Z. A. Dwiyantri and C. Prianto, "Prediksi Cuaca Kota Jakarta Menggunakan Metode Random Forest," *J. Tekno Insentif*, vol. 17, no. 2, pp. 127–137, 2023, doi: 10.36787/jti.v17i2.1136.
- [3] G. A. Mursianto, I. M. Falih, M. Irfan, T. Sakinah, and D. S. Prasvita, "Perbandingan Metode Klasifikasi Random Forest dan XGBoost Serta Implementasi Teknik SMOTE pada Kasus Prediksi Hujan," *J. Senamika*, vol. 2, no. 2, pp. 41–50, 2021.
- [4] A. Hot Iman, F. Ready Permana, G. Putro Wardana, R. Kemmy Rachmansyah, and M. Mega Santoni, "Perbandingan Algoritma Klasifikasi Random Forest dan Extreme Gradient Boosting pada Dataset Cuaca Provinsi DKI Jakarta Tahun 2018," *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, pp. 593–601, 2022, [Online]. Available: <https://katalog.data.go.id/dataset/data-prakiraan-cuaca-wilayah-provinsi-dki-jakarta-tahun-2018>.
- [5] R. Risanti, "Analisis Model Prediksi Cuaca Menggunakan Support Vector Machine, Gradient Boosting, Random Forest, Dan Decision Tree," vol. XII, pp. 119–128, 2024, doi: 10.21009/03.1201.fa18.
- [6] H. Nisya, C. Setianingsih, and W. Harjupa, "Prediksi Curah Hujan Dari Data Satelit Himawari-8 Menggunakan Metode K-Nearest Neighbor (KNN)," *e-Proceeding Eng.*, vol. 10, no. 1, pp. 729–735, 2023, [Online]. Available: <https://sharaku.eorc.jaxa.jp/GSMaP/>.
- [7] P. Alkhairi and A. P. Windarto, "Penerapan K-Means Cluster Pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara," pp. 762–767, 2019.
- [8] A. F. Budiantara and C. Budihartanti, "IMPLEMENTASI DATA MINING DALAM MANAJEMEN INVENTORY PADA PT . MASTERSYSTEM INFOTAMA MENGGUNAKAN METODE ALGORITMA APRIORI," vol. 7, no. 1, 2020.
- [9] A. M. Siregar, "Klasifikasi Untuk Prediksi Cuaca Menggunakan Esemble Learning," *Petir*, vol. 13, no. 2, pp. 138–147, 2020, doi: 10.33322/petir.v13i2.998.
- [10] Ashari Rakhmat Galih and Mutohar Wisnu, "Prakiraan Hujan menggunakan Metode RandomForest dan Cross Validation," *J. MIND J. / ISSN*, vol. 8, no. 2, pp. 173–187, 2023, [Online]. Available: <https://doi.org/10.26760/mindjournal.v8i2.173-187>.
- [11] K. S. Ningsih, N. J. Aruan, and A. T. A. A. Siahaan, "Aplikasi Buku Tamu Menggunakan Fitur Kamera Dan Ajax Berbasis Website Pada Kantor Dispora Kota Medan," *SITek J. Sains, Inform. dan Teknologi*, vol. 1, pp.

94–99, 2022.

- [12] H. A. Fauzi, K. A. Putra, and A. Tri, “Analisis Perbandingan Performa Web Service Menggunakan Bahasa Pemrograman Python , Php ,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 1, pp. 237–245, 2018, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/781>.

DAFTAR GAMBAR

Gambar 1 Struktur umum model <i>Random Forest</i>	11
Gambar 2 Tahapan Metode Penelitian.....	12