

Algoritma Klasifikasi Multilayer Perceptron Dalam Analisa Data Kebakaran Hutan

Haryani¹, Cucu Ika Agustyaningrum², Artika Surniandari³, Sucitra Sahara⁴, Ratna Kurnia Sari⁵

^{1,2,3,4,5} Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika

e-mail: ¹haryani.hyi@bsi.ac.id, ^{2*}cucu.cuk@bsi.ac.id, ³artika.ats@bsi.ac.id, ⁴sucitra.scr@bsi.ac.id, ⁵ratna.rus@bsi.ac.id

Diterima	Direvisi	Disetujui
05-05-2023	08-05-2023	31-05-2023

Abstrak - Kebakaran hutan atau yang sering disebut dengan wildfire merupakan salah satu isu lingkungan yang utama karena berdampak negatif terhadap kelestarian hutan, merugikan lingkungan dan ekonomi, serta merugikan masyarakat. Kebakaran hutan adalah kondisi di mana hutan terbakar, merusak hasil hutan dan menyebabkan kerusakan ekologi dan ekonomi. Tujuan dari peramalan kebakaran hutan adalah untuk mengetahui seberapa sering terjadi kebakaran hutan. Oleh karena itu, proses analisis data dilakukan dengan menggunakan teknik machine learning tradisional melalui metode Random Forest, Decision Tree, Logistic Regression, Naive Bayes dan Multilayer Perceptron. Mengetahui keakuratan dan nilai hasil F1 memungkinkan membandingkan metode ini dengan bahasa pemrograman Python. Hasil pengujian menunjukkan bahwa pendekatan Multilayer Perceptron mengungguli metode Random Forest, Decision Tree, Logistic Regression dan Nave Bayes dengan nilai akurasi masing-masing sebesar 93,35% dan F1 Score 93,69% dengan ukuran hidden layer sebesar 64,64. Dibandingkan dengan pendekatan lain yang dipelajari, nilai metode multilayer perceptron cukup signifikan. Penelitian ini dapat membantu menentukan kemungkinan kebakaran hutan.

Kata Kunci: Kebakaran Hutan, Metode, Multilayer Perceptron

Abstract - Forest fires or what is often referred to as forest fires is one of the main environmental issues because it has a negative impact on forest sustainability, harms the environment and the economy, and harms society. Forest fires are conditions in which forests burn, destroying forest products and causing ecological and economic damage. The purpose of forest forest fires is to find out how often forest fires occur. Therefore, the process of data analysis is carried out using traditional machine learning techniques through the Random Forest, Decision Tree, Logistic Regression, Naive Bayes and Multilayer Perceptron methods. Knowing the accuracy and value of the F1 result makes it possible to compare this method with the Python programming language. The test results show that the Multilayer Perceptron approach outperforms the Random Forest, Decision Tree, Logistic Regression and Nave Bayes methods with respective accuracy values of 93.35% and F1 Score of 93.69% with a hidden layer size of 64.64. Compared to other approaches studied, the value of the multilayer perceptron method is quite significant. This research can help determine the likelihood of forest fires.

Keywords: Forest Fire, Method, Multilayer Perceptron

PENDAHULUAN

Setiap bangsa di muka bumi membutuhkan hutan dan ekosistem hutan untuk bertahan hidup dan berkembang secara sosial, ekonomi dan lingkungan. Hutan dianggap secara permanen dan serius terancam oleh kebakaran hutan. Khusus untuk kebakaran besar, efek merugikan dari kebakaran hutan bertahan selama puluhan tahun setelah terbakar. Pencegahan kebakaran hutan adalah salah satu masalah yang paling krusial, dan ini melibatkan berbagai tindakan proaktif selain pembalasan seperti pemadaman kebakaran (Baranovskiy & Zharikova, 2014).

Kebakaran hutan atau yang sering disebut dengan wildfire merupakan salah satu isu lingkungan

yang utama karena berdampak negatif terhadap kelestarian hutan, merugikan lingkungan dan ekonomi, serta merugikan masyarakat. Jutaan hektar (ha) hutan di seluruh dunia mengalami kerusakan setiap tahun akibat kejadian tersebut, yang disebabkan oleh berbagai sumber (termasuk kelalaian manusia dan sambaran petir) (Cortez & Morais, 2007).

di kawasan hutan, yang membakar jutaan hektar setiap tahun dan bertanggung jawab atas hilangnya keanekaragaman hayati, kualitas tanah, dan penangkapan CO₂. Kerentanan hutan dan kawasan sekitarnya, yaitu permukiman manusia dan infrastruktur, terhadap kebakaran merupakan perhatian utama bagi masyarakat di banyak ekosistem

darat dunia. Meningkatnya perubahan dalam proses sosio-ekonomi dan iklim yang mengarah pada modifikasi ekstensif terhadap lingkungan alam dan periode kekeringan yang berkepanjangan telah menempatkan tuntutan yang kuat pada pihak berwenang dan pembuat keputusan untuk membatasi kawasan hutan secara temporal dan spasial dalam hal kerentanan terhadap kebakaran. Mengidentifikasi daerah dengan kerentanan kebakaran yang tinggi atau sangat tinggi adalah suatu keharusan agar berhasil merancang rencana pengelolaan kebakaran dan mengalokasikan sumber daya pemadam kebakaran. Untuk tujuan ini, pendekatan dan alat yang kuat diperlukan untuk memungkinkan manajer dan insinyur memprediksi secara akurat waktu, lokasi, dan tingkat kebakaran di masa mendatang. Perbaikan dalam teknik untuk memprediksi kerentanan kebakaran dan menggambarkan kawasan hutan menurut tingkat kerentanan yang berbeda dapat membantu pengelola hutan dan pembuat kebijakan mencapai pemahaman yang lebih baik tentang kebakaran, yang memfasilitasi pengembangan langkah-langkah pencegahan untuk hutan rawan kebakaran (Pham et al., 2020).

Kebakaran hutan adalah suatu keadaan dimana hutan dilalap api sehingga menimbulkan kerusakan hasil hutan yang mengakibatkan kerugian ekonomi dan lingkungan. Kebakaran terbesar, yang masing-masing menghabiskan lebih dari 2,6 juta dan 1,6 juta hektar hutan dan lahan Indonesia, terjadi di Indonesia selama periode lima tahun 2015–2019. Menurut data Kementerian Lingkungan Hidup, musim kemarau panjang di Indonesia dan kenaikan permukaan air laut berkontribusi pada dua kebakaran besar yang terjadi sekitar tahun itu. Sekitar 29% dari dua kebakaran tersebut berlokasi di lahan gambut.

Berbagai penelitian tentang kebakaran hutan di Indonesia yang pernah terjadi telah dilakukan, antara lain karya (Kresimo Negoro et al., 2022) “Analisis Kebakaran Hutan dan Lokasi di Provinsi Riau Menggunakan Metode C4.5” adalah judul penelitian tersebut. Karena keunggulannya yang signifikan dibandingkan algoritma pesaing, metode klasifikasi pohon keputusan C4.5 digunakan secara luas (Manalu et al., 2020). Berdasarkan temuan penelitian ini, kebakaran hutan dan lahan di Provinsi Riau sangat dipengaruhi oleh kondisi lingkungan seperti kelembaban, cuaca, dan kecepatan angin. Hasil analisis kelembaban rendah (kering) dan cuaca cerah dengan kecepatan angin tinggi dapat menunjukkan peluang terjadinya kebakaran hutan lebih tinggi. Sedangkan cuaca mendung dan kecepatan angin kencang dapat menyebabkan kebakaran hutan, meskipun persentasenya lebih kecil. Metode algoritma C.45 digunakan dalam penelitian ini dengan uji persamaan, dan hasilnya dibandingkan dengan matriks konfusi.

Pada penelitian lain oleh (Husen et al., 2022) berjudul “Analisis Prakiraan Kebakaran Hutan” dengan menggunakan Algoritma Random Forest

Classifier, penelitian ini mengembangkan konsep sistem prakiraan kebakaran hutan yang menjadi salah satu acuan kebijakan pemerintah dalam menentukan tindakan pencegahan. Studi ini memodelkan model algoritma random forest menggunakan data kebakaran hutan tahunan di Indonesia dengan harapan dapat membantu pemerintah dalam mencegah kebakaran hutan melalui kebijakan hukumnya dan analisis yang ada dapat digunakan dengan Teknologi Modifikasi Cuaca Center (BBTMC), yang dapat membantu menentukan kapan perubahan cuaca dapat dilakukan.

Kajian oleh (Yandi et al., 2021) berjudul “Prediksi Titik Panas Kebakaran Hutan Menggunakan Model Regresi SVM (Support Vector Machine) Pada Data Kebakaran Hutan Daops Manggala Agni Oki Provinsi Sumatera Selatan Tahun 2019” Metode prediksi data yang digunakan adalah SVM (Support Vector Machine) Metode (algoritma regresi mesin) dengan informasi (tanggal, waktu, satelit, akurasi, area, area sebagian, kelembaban dan suhu). Hasil penelitian menunjukkan hasil analisis yang cukup baik, nilai RSME sebesar 2,1 dan nilai R2 sebesar 0,83. dimana pada tahun 2021 sebagian besar hotspot akan muncul dari proses tersebut. Sedangkan untuk data tahun 2022, Kecamatan Cegal memiliki jumlah titik api terbanyak yaitu 571 titik api. Dan untuk meningkatkan visualisasi data, hasil prediksi hotspot divisualisasikan dalam bentuk heatmap.

Selain itu, (Ayuningtyas & Prasetyo, 2020) melakukan penelitian dengan judul Pemanfaatan Teknologi Machine Learning Untuk Klasifikasi Daerah Prene Kekeringan di Daerah Istimewa Yogyakarta menggunakan citra Landsat 8 Operational Land Imager (OLI). Penelitian ini dilakukan untuk memprediksi area yang berisiko dari citra satelit menggunakan metode autokorelasi pembelajaran mesin dan jaringan syaraf tiruan. Hasil analisis Moran's I menunjukkan bahwa semua indeks vegetasi dan prediksi JST memiliki autokoreksi positif.

Penelitian yang dilakukan oleh (Pratiwi et al., 2021) “Klasifikasi Kebakaran Hutan dan Kebakaran Lahan Menggunakan Algoritma Naive Bayes” Penelitian ini menggunakan data kebakaran hutan di Kabupaten Pelalawan tahun 2015 sampai dengan tahun 2019 dengan menggunakan metode Naive Bayes. Titik panas yang akan dianalisa terdiri dari suhu, kelembaban, curah hujan, kecepatan angin dan kategori. Metode klasifikasi berdasarkan algoritma Naive Bayes dapat digunakan untuk pencegahan sebelum kebakaran hutan dan lahan berkembang.

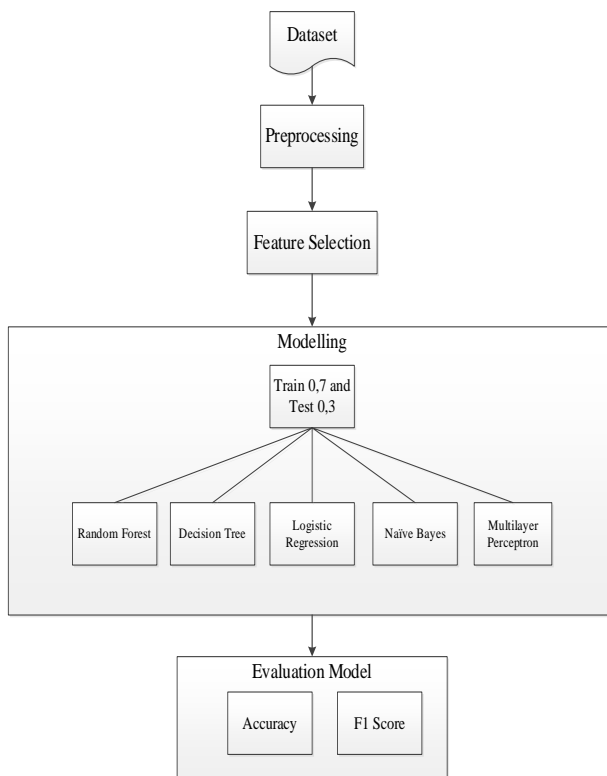
Diantara beberapa penelitian terkait, penulis membandingkan lima algoritma dalam penelitian ini yaitu Random Forest, Decision Tree, Logistic Regression, Naive Bayes dan Multilayer Perceptron. Metode prediksi yang sering digunakan dalam penelitian adalah MLP (Multilayer Perceptron Neural Network) (Manalu et al., 2020). Ciri MLP adalah keunggulannya dalam menentukan nilai bobot yang

lebih baik dibandingkan dengan metode lainnya; MLP dapat digunakan tanpa uang muka Informasi; Algoritma dapat dengan mudah diimplementasikan; dan dapat memecahkan masalah linear dan non-linear (Manalu et al., 2020). Model algoritma dengan tingkat perbandingan tertinggi dipilih dari lima perbandingan. Berdasarkan hasil model yang diperoleh, diharapkan dapat membantu pemerintah Indonesia untuk melakukan tindakan pencegahan yang paling tepat di masa mendatang.

Maka dari itu, penelitian ini fokus pada kedalam size metode multilayer perceptron yang merupakan hasil akurasi yang tinggi dengan menggunakan bahasa pemrograman python, sehingga mendapatkan hasil akurat untuk menentukan prediksi kebakaran hutan.

METODE PENELITIAN

Penelitian ini akan dilakukan dalam beberapa langkah seperti yang ditunjukkan pada Gambar 1.



Gambar 1 Tahapan Penelitian

Model yang dikembangkan untuk prediksi kebakaran hutan melalui tahapan data, preprocessing, pemilihan fitur dan evaluasi model teknik penelitian. mulai mengumpulkan data melalui UCI Machine Learning Repository. Setelah data diubah menjadi data mentah untuk pra-pemrosesan, fitur dipilih menggunakan Python, dan pemilihan fitur kemudian divalidasi dengan pengujian menggunakan algoritma pembelajaran mesin standar seperti Hutan Acak, Pohon Keputusan, Regresi Logistik, Naive Bayes dan Multilayer. metode Perceptron. Pendidikan.

Pada fase pemodelan, lima algoritma - Random Forest, Decision Tree, Logistic Regression, Naive Bayes dan Multilayer Perceptron - dibandingkan. Data yang digunakan diolah dengan uji latih dengan nilai latih 0,7 dan nilai uji 0,2 sebelum dilanjutkan dengan pemodelan. Langkah selanjutnya adalah menggunakan algoritme pembelajaran mesin standar untuk memeriksa data yang dihasilkan oleh bahasa pemrograman Python untuk menentukan teknik mana yang menghasilkan hasil terbaik untuk kumpulan data kebakaran. Ini dilakukan setelah preprocessing data, pemilihan fitur, pemodelan dan pengujian.

1. Penerapan Metode Penelitian

Data Set Kebakaran Hutan digunakan untuk menerapkan metodologi studi dalam lima tahap, yaitu:

Data Set

Untuk memperkirakan probabilitas kebakaran, 517 data set dengan 13 atribut dan 1 kelas dikumpulkan sebagai data sekunder dari UCI Machine Learning Repository. Masalah ini dapat diatasi dan diperoleh hasil yang lebih sederhana, cepat dan akurat dengan menggunakan teknik klasifikasi dan prediktabilitas serta akurasi yang maksimal. Dengan membandingkan Random Forest, Decision Tree, Logistic Regression, Nave Bayes dan Multilayer Perceptron menggunakan Python sebagai bahasa pemrograman studi, hasilnya dapat dibangun dengan cara yang sangat dapat diprediksi dan akurat.

Preprocessing

Untuk tahap penelitian ini terkumpul 517 data, total 13 atribut dan 1 kategori. Data ini dianalisis untuk memprediksi kebakaran hutan berdasarkan deskripsi fitur saat ini. Pemilihan data yang memerlukan penilaian terhadap kualitas jenis data yang akan diolah merupakan langkah awal dalam persiapan data.

Proses pembersihan data dilakukan setelah proses seleksi data selesai. Dalam prosedur ini, Anda harus mencoba menemukan nilai yang hilang.

Seleksi Fitur (Feature Selection)

Ini digunakan untuk mengidentifikasi fitur yang memiliki dampak terbesar pada data selama proses pemilihan. Pemodelan algoritma pembelajaran mesin tradisional kemudian menggunakan pelatihan dan pengujian untuk distribusi data.

Modeling

Proses prediksi dengan metode yang diusulkan melakukan langkah-langkah pemodelan dengan beberapa cara yang berbeda. Untuk mengevaluasi akurasi dan nilai f1 dalam prediksi kebakaran, algoritme pembelajaran mesin yang diusulkan menggunakan bahasa pemrograman Python dan mencakup Random Forest, Decision

Tree, regresi logistik, Nave Bayes, dan pendekatan multilayer perceptron (Agustyaningrum et al., 2021).

Machine Learning adalah pengenalan otomatis pola signifikan dalam data. Komputer dapat mempelajari berbagai hal dari manusia melalui pembelajaran mesin. Tanpa pemrograman khusus, komputer dapat belajar mengolah data yang disajikan kepadanya. Algoritma pembelajaran mesin digunakan untuk melatih komputer memproses data

Konsep Random Forest digunakan untuk menghasilkan sejumlah besar pohon keputusan berkorelasi, dan setiap pohon keputusan bertindak sebagai satu set model. Setiap pohon keputusan menentukan prediksi kelas dan keputusan akhir didasarkan pada kinerja maksimum (Kabir et al., 2019). Metode klasifikasi hutan acak didasarkan pada pendekatan pohon keputusan di mana atribut pada setiap node dipilih secara acak untuk menentukan klasifikasi. Suara yang dikembalikan oleh pohon keputusan terutama digunakan dalam klasifikasi data (Ratnawati & Sulistyanningrum, 2019). Menggunakan sistem voting (angka tertinggi) untuk menggabungkan pengklasifikasi independen (CARTS) dari distribusi yang sama, hutan acak menghasilkan prediksi klasifikasi. Mengurangi korelasi dapat mengurangi kesalahan prediksi hutan acak yang merupakan karakteristik dari hutan acak (As Sarofi et al., 2020). Rumus Random Forest sebagai berikut (Leonardo et al., 2020):

$$Entropy(Y) = - \sum_i P(Y) \log^2 p(Y) \quad (1)$$

$$= Entropy(Y) - \sum_v \epsilon_{values(a)} \frac{|Y_v|}{|Y_a|} Entropy(Y_v) \quad (2)$$

Keterangan :

Y = suite apartemen

P(c|Y) adalah rasio nilai di kelas Y dengan nilai di kelas C.

Nilai (a) = Nilai yang mungkin jika ditetapkan.

Y_v = subkelas Y dengan kelas v relatif terhadap kelas a.

Ya = Semua nilai yang cocok dengan a.

Support Vector Machine Teknik pembelajaran mesin yang disebut Support Vector Machine bekerja berdasarkan prinsip minimalisasi risiko struktural (Saputra et al., 2022). Karena memerlukan tujuan pembelajaran khusus selama pelatihan, maka Support Vector Machine (SVM) merupakan metode klasifikasi yang terintegrasi (diawasi)(Nurachim, 2019). Rumus metode Support Vector Machine Learning (Zulfikar & Lukman, 2016).

$$similarity = \frac{\sum_{i=1}^n f(T_i, S_i)}{w_i} \quad (3)$$

Keterangan:

T: Kasus baru

S: kasing dalam penyimpanan

n : jumlah atribut

I: atribut individu antara 1 dan n

f: Fungsi kesamaan TRIBUTE antara kasus T dan kasus S

W: bobot yang diberikan pada atribut ke-i.

Logistic Regression adalah bagian dari proses klasifikasi yang diawasi. Penggunaan algoritma ini telah meningkat pesat dalam beberapa tahun terakhir karena popularitasnya. Kurva ini sigmoidal. Itu termasuk dalam kategori regresi logistik. Untuk memahami pemaparan matematis dari penjelasannya, mari kita mulai dengan rumus regresi linier sederhana (Shah et al., 2020)

$$y = b_0 + b_1 * x \quad (4)$$

Jadi kami menemukan fungsi sigmoid dan hasilnya diberikan oleh rumus.

$$p = \frac{1}{1+e^{-y}} \quad (5)$$

Sekarang satu formula telah diganti dengan yang lain untuk mendapatkan nilai y, kami memiliki formula regresi logistik kami.

$$\text{logit}(S) = b_0 + b_1M_1 + b_2M_2 + b_3M_3 \dots bkMk \dots \quad (6)$$

di mana S berarti kemungkinan fitur yang menarik.

Nilai yang diprediksi adalah M₁, M₂, M₃, ... M_k.

Perpotongan model-model tersebut adalah b₀, b₁, b₂, b₃, ... b_k.

Multilayer Perceptron adalah jaringan saraf multilayer yang terdiri dari beberapa neuron yang saling berhubungan yang bobotnya terkait satu sama lain. Dengan lapisan masukan, satu atau lebih lapisan tersembunyi dan lapisan keluaran, neuron-neuron ini disusun menjadi beberapa lapisan (Irfan et al., 2017).

Naïve Bayes adalah algoritma yang mudah dipahami berdasarkan teorema probabilitas bersyarat Bayes. Hal ini dilakukan untuk mengklasifikasikan data berdasarkan berapa kali data deskriptor muncul dalam pelatihan. Algoritma Naive Bayes mengasumsikan bahwa semua data sama-sama independen. Sebuah metode untuk menemukan ketergantungan antar fitur dalam set pelatihan menggunakan asumsi ini (Agustyaningrum et al., 2020).

Evaluasi

Proses prediksi dilakukan menggunakan algoritma pembelajaran mesin tradisional dalam fase estimasi untuk memeriksa akurasi dan F1 Score dalam tingkat keberhasilan dan kegagalan kebakaran hutan. Metode tersebut antara lain Random Forest, Decision Tree, Logistic Regression, Naive Bayes dan Multilayer Perceptron.

2. Metode Pengumpulan Data

Data primer dan data sekunder merupakan dua kategori yang teknik pengumpulan datanya yang dapat dibedakan. Sedangkan data sekunder diperoleh dari peneliti yang melakukan penelitian yang relevan, sedangkan data primer dikumpulkan langsung dari sumbernya. Pada Tabel 1 dan 2, total 517 dataset dengan 13 atribut dan 1 atribut kelas dari dataset kebakaran hutan diambil dari UCI Machine Learning Repository untuk digunakan dalam penelitian ini.

Tabel 1 Deskripsi atribut dataset survival kebakaran hutan

attribute name	data type	description
X	Numerik	x-axis spatial coordinate within the Montesinho park map: 1 to 9
Y	Numerik	y-axis spatial coordinate within the Montesinho park map: 2 to 9
month	Kategori	month of the year: 'jan' to 'dec'
day	Kategori	day of the week: 'mon' to 'sun'
FFMC	Numerik	FFMC index from the FWI system: 18.7 to 96.20
DMC	Numerik	DMC index from the FWI system: 1.1 to 291.3
DC	Numerik	DC index from the FWI system: 7.9 to 860.6
ISI	Numerik	ISI index from the FWI system: 0.0 to 56.10
temp	Numerik	temperature in Celsius degrees: 2.2 to 33.30
RH	Numerik	relative humidity in %: 15.0 to 100
wind	Numerik	wind speed in km/h: 0.40 to 9.40
rain	Numerik	outside rain in mm/m2 : 0.0 to 6.4
area	Numerik	the burned area of the forest (in ha): 0.00 to 1090.84

Tabel 2 Atribut Numerik dan Kategori Analisis Perilaku Pengguna

Attribute Name	Min. Value	Max. Value	STD
X	1	9	2.31
Y	2	9	1.23
FFMC	18.70	96.20	5.52
DMC	1.10	291.30	64.05
DC	7.90	860.6	248.07
ISI	0	56.1	4.56
temp	2.20	33.3	5.81
RH	15	100	16.32
wind	0.4	9.4	1.79
rain	0	6.4	0.3
area	0	1090.84	63.66

HASIL DAN PEMBAHASAN

Dengan total 517 dataset, 13 atribut dan satu atribut kelas, data sekunder yang dibutuhkan untuk prediksi kebakaran dikumpulkan dari UCI Machine Learning Repository. Menurut penelitian "A Data Mining Approach to Predict Forest Fire using Meteorological Data" oleh Paulo Cortez dan Anibal Morais dengan menggunakan metode SVM dan Random Forest, suhu 9,95, kelembaban relatif 0,56,

angin 0,64 dan hujan juga 2: 45. memberikan suhu relatif 73,2%, kelembaban relatif 4,1%, angin 4,7%, dan curah hujan 18%, memberikan pola varians SVM terbaik.

Hasil penelitian terdiri dari data kualitatif dan kuantitatif yang dihitung dengan menggunakan model yang diusulkan. Semua dataset yang tersedia digunakan dalam penelitian ini. Eksperimen dan pengujian penelitian dilakukan dengan memproyeksikan materi menggunakan metode pembelajaran mesin tradisional. Eksperimen ini dilakukan pada dataset yang diterima berdasarkan hasil preprocessing, pemilihan fitur, pemodelan dan penilaian yang dilakukan dengan menggunakan bahasa pemrograman Python dan Google Collaboratory.

1. Preprocessing, Step Validation dan Algoritma Conventional Machine Learning

Tahap *Preprocessing* dimana proses pemilihan data yang memerlukan penilaian terhadap kualitas jenis data yang akan diolah, mulai dari atribut dengan data type kategori dirubah menjadi numerik. Dengan ini atribut month dan day dengan tipe data awal kategori berubah menjadi numerik

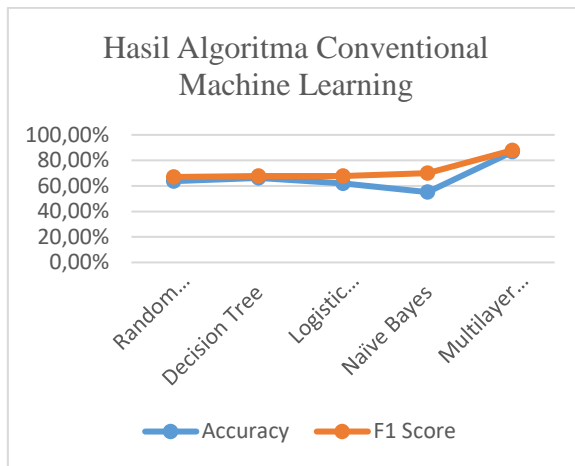
Pada Step Validasi atau biasa yang disebut dengan feature selection (seleksi fitur) digunakan untuk mengidentifikasi fitur yang memiliki dampak terbesar pada data selama proses pemilihan, dalam penelitian prediksi kebakaran hutan ini seleksi fitur yang digunakan dimana menghapus beberapa fitur yang tidak digunakan atau yang kurang berpengaruh dalam proses pengelolaan, sehingga yang diambil hanya fitur yang berpengaruh akan hasil keluaran prediksi kebaran hutan.

Nilai-nilai berikut ditentukan dalam studi yang menggunakan data preprocessing validasi data prediksi kebakaran hutan:

Tabel 3 Hasil Perbandingan Nilai Prediksi Kebakaran Hutan

Model	Accuracy	F1 Score
Random Forest	63,71%	67%
Decision Tree	66,2%	67,73%
Logistic Regression	62,05%	67,61%
Naïve Bayes	55,13%	69,89%
Multilayer Percepton	93,35%	93,69%

Metode Multilayer Peceptron memiliki nilai akurasi 93,35% dan Skor F1 93,69% lebih tinggi dibandingkan dengan metode Random Forest, Decision Tree, Logistic Regression dan Nave Bayes, menurut hasil analisis data prediksi kebakaran hutan menggunakan algoritma machine learning tradisional. Gambar 2 mengilustrasikan hal ini dengan perbedaan nilai 25% sampai 28%.



Gambar 2 Hasil Nilai Algoritma Machine Learning Konvensional

Pengujian dijalankan dengan mengoptimalkan algoritma conventional machine learning menggunakan pendekatan perceptron multilayer, yang memiliki nilai lebih tinggi daripada metode lain dan penelitian terkait, dibuktikan dengan menganalisis confusion matrix. Multilayer Perceptron menghasilkan akurasi 93,35% dan skor F1 93,69% merupakan hasil lebih besar daripada metode Random Forest, Decision Tree, Logistic Regression, dan Nave Bayes. Perbedaan akurasi rata-rata adalah 14,03%, dan skor F1 adalah 5,29%, menurut angka tersebut.

KESIMPULAN

Prosedur preprocessing untuk studi data prediksi kebakaran hutan yang diperoleh menggunakan data selection, data cleaning dan feature selection. Algoritma Conventional Machine Learning untuk data mining dapat menangani data dengan baik, dengan menggunakan metode multilayer perceptron dalam parameter train 0,7 dan parameter test 0,3. Metode multilayer perceptron menghasilkan akurasi 93,35% dan Skor F1 93,69% dengan ukuran lapisan tersembunyi 64,64, mengungguli metode Random Forest, Decision Tree, Logistic Regression, dan Nave Bayes. Nilai ini cukup dominan dibandingkan dengan metode lainnya. Kajian ini dapat menentukan peluang terjadinya kebakaran hutan dan diharapkan pada penelitian selanjutnya dapat dikembangkan dengan menggali ukuran hidden layer untuk pelaporan kebakaran hutan yang lebih akurat.

REFERENSI

Agustyaningrum, C. I., Gata, W., Nurfalih, R., & Radiyah, U. (2020). Komparasi Algoritma Naive Bayes, Random Forest Dan Svm Untuk Memprediksi Niat. *Jurnal Informatika*, 20(2).
Agustyaningrum, C. I., Haris, M., Aryanti, R., & Misriati, T. (2021). Online Shopper Intention

Analysis Using Conventional Machine Learning And Deep Neural Network Classification Algorithm. *Jurnal Penelitian Pos Dan Informatika*, 11(1), 89–100. <https://doi.org/10.17933/jppi.v11i1.341>
As Sarofi, M. A., Irhamah, I., & Mukarromah, A. (2020). Identifikasi Genre Musik dengan Menggunakan Metode Random Forest. *Jurnal Sains Dan Seni ITS*, 9(1), 79–86. <https://doi.org/10.12962/j23373520.v9i1.51311>
Ayuningtyas, F., & Prasetyo, S. Y. J. (2020). Pemanfaatan Teknologi Machine Learning Untuk Klasifikasi Wilayah Risiko Kekeringan di Daerah Istimewa Yogyakarta Menggunakan Citra Landsat 8 Operational Land Imager (OLI). *Jurnal Transformatika*, 18(1), 13. <https://doi.org/10.26623/transformatika.v18i1.2140>
Baranovskiy, N., & Zharikova, M. (2014). A web-oriented geoinformation system application for forest fire danger prediction in typical forests of the Ukraine. *Lecture Notes in Geoinformation and Cartography*, 0(199669), 13–22. https://doi.org/10.1007/978-3-319-08180-9_2
Cortez, P., & Morais, A. (2007). A Data Mining Approach to Predict Forest Fires using Meteorological Data. *Proceedings of 13th Portugese Conference on Artificial Intelligence*, 512–523. <http://www.dsi.uminho.pt/~pcortez/fires.pdf>
Husen, D., Sandi, D., & Bumbungan, S. (2022). Analisis Prediksi Kebakaran Hutan dengan Menggunakan Algoritma Random Forest Classifier Kebakaran hutan dan lahan di Indonesia telah menjadi perhatian dunia internasional khususnya sejak kebakaran hutan yang terjadi pada tahun 80-an [2]. *Penyebab kebaka*. 16, 150–155.
Irfan, M., Ardi Sumbodo, B. A., & Candradewi, I. (2017). Sistem Klasifikasi Kendaraan Berbasis Pengolahan Citra Digital dengan Metode Multilayer Perceptron. *IJEIS (Indonesian Journal of Electronics and Instrumentation Systems)*, 7(2), 139. <https://doi.org/10.22146/ijeis.18260>
Kabir, M. R., Ashraf, F. Bin, & Ajwad, R. (2019). Analysis of different predicting model for online shoppers' purchase intention from empirical data. *2019 22nd International Conference on Computer and Information Technology, ICCIT 2019, March 2020*. <https://doi.org/10.1109/ICCIT48885.2019.9038521>
Kresimo Negoro, N., diana, M., Izul Ula, M., & Dwi Insani, F. (2022). Analisis Kebakaran pada Hutan dan Lokasi Lahan di Provinsi Riau Menggunakan Metode C4.5. *Maret*, 7(1), 107–114.
Leonardo, R., Pratama, J., & Chrisnatalis, C. (2020).

- Perbandingan Metode Random Forest Dan Naïve Bayes Dalam Prediksi Keberhasilan Klien Telemarketing. *Jurnal Teknologi Dan Ilmu Komputer Prima (Jutikomp)*, 3(2), 1–5.
- Manalu, D. R., Zarlis, M., Mawengkang, H., & Sitompul, O. S. (2020). *Forest Fire Prediction in Northern Sumatera using Support Vector Machine Based on the Fire Weather Index. April 2021*, 187–196. <https://doi.org/10.5121/csit.2020.101915>
- Nurachim, R. I. (2019). *Pemilihan Model Prediksi Indeks Harga Saham Yang Dikembangkan Berdasarkan Algoritma Support Vector Machine (Svm) Atau Multilayer Perceptron (Mlp) Studi Kasus : Saham Pt Telekomunikasi Indonesia Tbk*. 5(1), 29–35.
- Pham, B. T., Jaafari, A., Avand, M., Al-Ansari, N., Du, T. D., Hai Yen, H. P., Phong, T. Van, Nguyen, D. H., Van Le, H., Mafi-Gholami, D., Prakash, I., Thuy, H. T., & Tuyen, T. T. (2020). Performance evaluation of machine learning methods for forest fire modeling and prediction. *Symmetry*, 12(6), 1–21. <https://doi.org/10.3390/SYM12061022>
- Pratiwi, T. A., Irsyad, M., & Kurniawan, R. (2021). Klasifikasi Kebakaran Hutan dan Lahan Menggunakan Algoritma Naïve Bayes (Studi Kasus: Provinsi Riau). *Jurnal Sistem Dan Teknologi Informasi (Justin)*, 9(2), 101. <https://doi.org/10.26418/justin.v9i2.42823>
- Ratnawati, L., & Sulistyaningrum, D. R. (2019). Penerapan Random Forest untuk Mengukur Tingkat Keparahan Penyakit. *Jurnal Sains Dan Seni Its*, 8(2), A71–A77.
- Saputra, R. A., Puspitasari, D., & Baidawi, T. (2022). *Deteksi Kematangan Buah Melon dengan Algoritma Support Vector Machine Berbasis Ekstraksi Fitur GLCM*. 4(2).
- Shah, K., Patel, H., Sanghvi, D., & Shah, M. (2020). A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification. *Augmented Human Research*, 5(1). <https://doi.org/10.1007/s41133-020-00032-0>
- Yandi, J., Kurniawan, T. B., Negara, E. S., & Akbar, M. (2021). Prediksi Lokasi Titik Panas Kebaran Hutan menggunakan Model Regresion SVM (Support Vector Machine) pada Data Kebakaran Hutan Daops Manggala Agni Oki Provinsi Sumatera Selatan Tahun 2019. *InfoTekJar : Jurnal Nasional Informatika Dan Teknologi Jaringan*, 6(1), 10–15.
- Zulfikar, W. B., & Lukman, N. (2016). Perbandingan Naive Bayes Classifier Dengan Nearest Neighbor Untuk Identifikasi Penyakit Mata. *Jurnal Online Informatika*, 1(2), 82–86. <https://doi.org/10.15575/join.v1i2.33>