



Analisa Komparasi Algoritma *Naïve Bayes*, *Decision Tree* Dan KKN Untuk Klasifikasi Kebakaran Hutan Pada Wilayah Aljazair

Muhammad Fadhiil Alamsyah¹, Tri Putra Satriawan², Femmy Novica Ramadanis³, Rahma Anugrah Mulyawan⁴, Candra Edmond⁵, Ricky Firmansyah⁶
1,2,3,4,5,6 ARS University

Alamat: Jl.Sekolah International No 1-4 Bandung

Korespondensi Email : muhammadfadhilalamsyah@gmail.com¹, tripsatriawan30@gmail.com²,
femmynovica097@gmail.com³, rahmamulyawan121003@gmail.com⁴, candrasafra0@gmail.com⁵,
ricky@ars.ac.id⁶

ABSTRACT. *The Mediterranean region, in particular Algeria, is experiencing serious challenges due to the increased opportunities for forest fires. Since the mid-1970s, there has been a 50% reduction in rainfall over northwestern Algeria, making northern Algeria particularly vulnerable to the problem for many years. More than 37,000 hectares of sensitive forest are lost every year due to this extreme drought. The findings of this study, which assessed the hazard of forest fires from 2006 to 2019, agree with those of Bentchakal, Chibane (2022), who examined the problems caused by forest fires in the region. The aim of this investigation is to gain a better understanding of the problems caused by local forest fires and to use that expertise to provide insight for the authors and readers of this report. The report was written by presenting the findings of observations made using the Rapid Miner classification approach, which includes the categorization of areas affected by forest fires. Data is collected using a variety of algorithmic techniques, including Naïve Bayes, KNN, and decision trees, which are used as tests of data to identify the most accurate results. The findings show that the Decision Tree technique has the best accuracy of 86.49% and provides a thorough explanation of the data.*

Keywords: *Algeria, Decision trees, KNN, Forest fires, Naïve bayes.*

ABSTRAK. Kawasan Mediterania, khususnya Aljazair, mengalami tantangan berat akibat meningkatnya peluang terjadinya kebakaran hutan. Sejak pertengahan tahun 1970-an, terjadi penurunan curah hujan sebesar 50% di wilayah barat laut Aljazair, membuat Aljazair utara sangat rentan terhadap masalah ini selama bertahun-tahun. Lebih dari 37.000 hektar hutan sensitif telah hilang setiap tahun akibat kekeringan ekstrem ini. Temuan penelitian ini, yang menilai bahaya kebakaran hutan dari tahun 2006 hingga 2019, sesuai dengan penelitian Bentchakal, Chibane (2022), yang meneliti masalah yang disebabkan oleh kebakaran hutan di wilayah tersebut. Tujuan investigasi ini adalah untuk mendapatkan pemahaman yang lebih baik tentang masalah yang disebabkan oleh kebakaran hutan lokal dan menggunakan keahlian tersebut untuk memberikan wawasan bagi penulis dan pembaca laporan ini. Laporan tersebut ditulis dengan menyajikan temuan pengamatan yang dilakukan dengan menggunakan pendekatan klasifikasi Rapid Miner, yang meliputi pengkategorian wilayah yang terkena dampak kebakaran hutan. Data dikumpulkan menggunakan berbagai teknik algoritme, termasuk Naive Bayes, KNN, dan pohon keputusan, yang digunakan sebagai pengujian data untuk mengidentifikasi hasil yang paling akurat. Temuan

menunjukkan bahwa teknik Decision Tree memiliki akurasi terbaik sebesar 86,49% dan memberikan penjelasan data yang menyeluruh.

Kata kunci : Aljazair, Decision tree, KNN, Kebakaran hutan, Naïve bayes.

LATAR BELAKANG

MASALAH SECARA UMUM

Sangat penting untuk memahami tingkat keparahan dan lokasi kebakaran hutan untuk mengambil tindakan yang tepat sambil membandingkan algoritma Naive Bayes, Decision Tree, dan KNN untuk mengklasifikasikan kebakaran hutan di wilayah Aljazair (Khedr, 2019). Data dikategorikan menggunakan algoritma klasifikasi, yang juga digunakan untuk meramalkan tingkat keparahan kebakaran hutan berdasarkan data historis. Tiga algoritma klasifikasi terkemuka yang dapat digunakan untuk masalah ini adalah Naive Bayes, Decision Trees, dan KNN (Mishra, 2020).

Laporan tersebut ditulis dengan menyajikan temuan pengamatan yang dilakukan dengan menggunakan pendekatan klasifikasi Rapid Miner, yang meliputi pengkategorian wilayah yang terkena dampak kebakaran hutan. Data dikumpulkan menggunakan berbagai teknik komputasi, termasuk Naive Bayes, pohon keputusan, dan KNN (Hussaini, 2021).

MASALAH SECARA KHUSUS

Secara global, kebakaran hutan menjadi perhatian yang merusak ekosistem dan mengakibatkan kerugian finansial yang signifikan. Analisis data sangat penting untuk mencegah dan memantau kebakaran hutan (Zargayouna 2018). Klasifikasi merupakan salah satu pendekatan yang dapat digunakan. Ada beberapa metode klasifikasi yang dapat digunakan, antara lain Naive Bayes, Decision Trees, dan K-Nearest Neighbor (KNN).yang dilakukan dengan menggunakan dataset algeria fires forest sebagai dataset yang akan dilakukan penelitian lebih lanjut menggunakan machine learning (Bista, 2018).

RUMUSAN MASALAH

- Rumusan masalah pada penelitian ini adalah:
 1. Bagaimana menetapkan performa Klasifikasi yang optimal dari komparasi algoritma naïve bayes ,Decision tree dan K-nn ?
 2. Bagaimana membandingkan hasil dari ketiga metode klasifikasi tersebut.setelah mengetahui data terakurat ?

- Tujuan Penelitian

1. Sesuai dengan rumusan masalah yang telah dibahas diatas maka tujuan dari penelitan ini adalah mengetahui algoritma naïve bayes ,Decision tree dan K-nn yang optimal untuk penelitian data kebakaran hutan algeria.
2. Dengan data tersebut kami ingin membandingkan metode mana yang lebih akurat antara algoritma decision tree , algoritma knn dan algoritma naïve naves untuk mengklasifikasikan dataset kebakaran hutan Algeria , dengan hasil tersebut membuktikan bahwa Decission tree lebih akurat untuk mengklasifikasikan data tersebut.

- Batasan Masalah

Adapun beberapa Batasan masalah dari penelitian ini ialah sebagai berikut :

1. Dataset yang digunakan dalam penelitian ini berasal dari data sekuler yang diperoleh dari UCI Repository berupa data kebakaran hutan Algeria
2. Algoritma Klasifikasi menggunakan metode naïve bayes dan decision tree dan KNN
3. Untuk metode perbandingan menggunakan metode naïve bayes dan decision tree. Dan KNN

- Penelitian Terkait

Penelitian serupa pernah dilakukan untuk menentukan perbandingan Algoritma Naive Bayes, Decision Tree, dan k-Nearest Neighbor untuk Klasifikasi Kebakaran Hutan di Aljazair merupakan salah satu contoh analisis terkait yang telah dilakukan sebelumnya pada dataset yang bersangkutan. Menurut hasil studi yang dilakukan pada tahun 2019 oleh Djahri, A., dan Boukhatem, tiga metode berbeda digunakan untuk menganalisis kualitas kategori data terkait pembakaran hutan. Menurut temuan, algoritma k-NN menawarkan klasifikasi ketajaman terbaik.yang dilakukan oleh Djahri pada tahun 2019 Sebuah studi perbandingan menggunakan algoritma pohon keputusan, Naive Bayes, dan KNN untuk mengkategorikan kebakaran hutan di Aljazair kemudian dilakukan, dengan hasil menunjukkan bahwa Pohon Keputusan, Naive Bayes, dan KNN melakukan yang terbaik secara keseluruhan. Pekerjaan ini dilakukan oleh Amara, H., dan Boukhatem pada tahun 2019. Menurut hasil, algoritma KNN memberikan akurasi klasifikasi yang sebaik mungkin dilansir oleh Amara pada tahun 2019 Dalam studi menyeluruh yang dilakukan oleh Arbi, M.B., dan Ouhbi,

A. , Evaluasi Naive Bayes, Pohon Keputusan, dan Algoritma KNN untuk Prediksi Kebakaran Hutan di Aljazair adalah subjek kalimat 2021. Dalam penelitian ini, yang mana diterbitkan dalam jurnal "Ilmu Terapan", efektivitas dan kinerja algoritma Naive Bayes, Decision Tree, dan KNN dievaluasi. Menurut temuannya, algoritma KNN memiliki akurasi dan kecepatan yang tinggi dalam menentukan kemungkinan terjadinya kebakaran hutan berdasarkan penelitian yang dilakukan oleh. Arbi pada tahun 2021

TINJAUAN PUSTAKA

1. Data Mining

Penambangan data adalah proses mengambil informasi berharga dari data besar dan menemukan pola yang sebelumnya tidak ada. Menemukan wawasan yang menarik dan praktis dari data yang ada adalah tujuannya. Untuk melakukan ini, pola dan hubungan dalam data dapat ditemukan menggunakan teknik statistik, pembelajaran mesin, dan pembelajaran mesin.(Han, 2018).

Banyak industri, termasuk bisnis, kesehatan, dan pemasaran, telah menggunakan penambangan data. Penambangan data digunakan dalam bisnis untuk memeriksa data pelanggan dan transaksional untuk memahami tren dan preferensi pelanggan.(Fayyad, 2018) Penambangan data digunakan dalam industri perawatan kesehatan untuk memeriksa data kesehatan pasien untuk memahami penyebab penyakit dan mengidentifikasi pasien yang membutuhkan terapi. Penambangan data adalah teknik yang digunakan dalam pemasaran untuk memeriksa data penjualan untuk lebih memahami bagaimana produk dan layanan dijual (Aggarwal, 2018).

2. Klasifikasi

Klasifikasi data mining adalah proses menambahkan data baru ke grup yang telah diidentifikasi berdasarkan karakteristik atau fitur dari data tersebut(Alpaydin, 2018). Hal ini mendorong mesin pembelajar untuk memasukkan data ke dalam pola dan menggunakan pola tersebut untuk menentukan data baru untuk kelas tersebut. Algoritma klasifikasi yang populer termasuk Naive Bayes, Decision Trees, dan Support Vector Machines (SVM) (Domingos, 2018).

Klasifikasi data mining digunakan di banyak bidang, termasuk keamanan, pemasaran, dan perawatan kesehatan. Klasifikasi data mining digunakan dalam pelayanan kesehatan untuk mengidentifikasi penyakit berdasarkan gejala pasien yang berpotensi berbahaya. Dalam beberapa kasus, penambangan data klasifikasi digunakan untuk memprediksi produk atau layanan yang akan

dibeli pelanggan berdasarkan preferensi pribadi mereka. Klasifikasi data mining digunakan untuk menentukan aktivitas ancaman berdasarkan aktivitas caman yang diketahui di dunia nyata.(Mitchell, 2018).

3. Decision tree

Decision Decision Tree adalah sistem klasifikasi untuk penambangan data yang menggunakan pohon keputusan untuk meramalkan kelas data baru.(Wang, 2019) Menggunakan atribut fitur, Decision Tree memisahkan data menjadi subset dan mendasarkan pilihan pada subset ini. Pohon keputusan ini akan terus mensegmentasi data hingga setiap subset data homogen (Zhang, 2019).

Lima tahun sebelumnya telah terlihat peningkatan signifikan dalam kualitas algoritme dan aplikasinya di berbagai industri. Beberapa penelitian berfokus pada perubahan metode pembuatan pohon untuk meningkatkan akurasi. Pohon keputusan juga telah digunakan dalam studi di industri seperti keamanan, pemasaran, dan perawatan Kesehatan (Li, 2019).

4. Naive Bayes

Algoritma klasifikasi penambangan data yang disebut Naive Bayes memanfaatkan teori probabilitas dan statistik untuk meramalkan kelas data baru. Landasan dari pendekatan ini adalah gagasan tentang independensi atribut, atau gagasan bahwa setiap fitur memiliki dampak yang berbeda pada kelas (Li, 2019).

Penelitian Penambangan Data Naive Bayes selama lima tahun sebelumnya telah berkonsentrasi pada peningkatan standar algoritme dan memperluas jangkauan aplikasi untuk algoritme tersebut. Dalam beberapa penelitian, akurasi ditingkatkan dengan mengubah fitur-fitur yang dianggap independen, menggunakan metode seperti jaringan Bayesian. Naive Bayes juga telah digunakan dalam sejumlah penelitian di industri seperti keamanan, pemasaran, dan perawatan kesehatan (Zhang, 2019).

5. Algoritma K-Nearest Neighbor (K-NN)

K-N K-Nearest Neighbor (K-NN) adalah teknik klasifikasi yang menentukan kelas objek berdasarkan data pelatihan yang dikumpulkan sebelumnya. Algoritma ini menentukan kelas item berdasarkan kelas data training yang jaraknya paling dekat dengan objek yang akan diprediksi dan

mengukur jarak antara keduanya. Nilai K dalam K-NN menunjukkan berapa banyak contoh pelatihan terdekat yang akan digunakan sebagai dasar untuk mengklasifikasikan objek. Algoritma k-NN digunakan di banyak aplikasi untuk menambang nilai data dan meningkatkan kegunaan layanan data karena mudah dan efektif untuk data dengan kelas yang terdefinisi dengan baik dan jarak terukur yang dapat diukur antar item (Kaur, 2021)

METODE

Pengujian yang melihat bagaimana data pelatihan dan uji didistribusikan ke kumpulan data yang besar. Data dari Repositori Pembelajaran Mesin UCI digunakan dalam kumpulan data saat ini. Kumpulan atau basis data sumber daya, area teoretis, dan generator data yang digunakan oleh komunitas pembelajaran mesin (pembelajaran mesin) dikenal sebagai Repositori Pembelajaran Mesin UCI (Lin 2020).

Komputer sekarang dapat diprogram oleh manusia sehingga mereka dapat mengidentifikasi pola atau mengambil informasi yang dimasukkan melalui pembelajaran mesin. Proses mengubah informasi menjadi pengetahuan adalah definisi belajar. Keluaran dari algoritma pembelajaran mesin adalah beberapa pengetahuan, yang biasanya berupa program komputer lain yang dapat menjalankan tugas, yang diwakili oleh data pelatihan, yang mencerminkan pengalaman. Dengan bantuan pembelajaran mesin, sistem klasifikasi data dapat mengidentifikasi pola dalam data dengan atau tanpa pelatihan. Dalam metode pembelajaran mesin, pengelompokan mengacu pada klasifikasi data (Mesran 2020).

Dataset akan digunakan dalam penelitian ini untuk menilai efektivitas strategi. Data yang digunakan terdiri dari dua dataset dari UCI Machine Learning Repository, salah satunya berisi informasi kebakaran hutan yang terjadi di Aljazair antara Juni 2012 dan September 2012. Dataset tersebut berisi 244 entri yang menggabungkan data dari dua lokasi Aljazair: Sidi Bel wilayah - abbes di barat laut Aljazair dan wilayah Bejaia di timur laut Aljazair. Ada 122 catatan untuk setiap daerah. 244 entri dalam dataset telah dikategorikan menjadi "terbakar" (138 kelas) dan "tidak terbakar" menggunakan 11 atribut dan 1 atribut output (106 kelas). Dalam hal ini, peneliti menggunakan aplikasi Rapidminer dengan metode klasifikasi dengan memakai algoritma Decision Tree, Naïve bayes, dan k-NN. Ketiga algoritma ini digunakan dalam pengklasifikasian data kebakaran hutan di Aljazair. Hasil dari penggunaan ketiga algoritma ini akan dibandingkan satu sama lain untuk mengetahui algoritma mana yang paling akurat untuk mengklasifikasi data kebakaran hutan Aljazair pada tahun 2012 – 2022 (Curt 2020).

- **Metode Kualitatif**

Metode pengembangan perangkat lunak kualitatif adalah salah satu pendekatan yang lebih fleksibel dan memfokuskan pada pemahaman secara mendalam tentang pengguna, lingkungan, dan konteks aplikasi yang akan dibuat. Metode ini menempatkan pengguna sebagai subjek utama dalam pengembangan perangkat lunak, dengan tujuan untuk memahami kebutuhan dan preferensi pengguna secara mendalam.

Menurut (McCusker, 2018). pemilihan penggunaan metode kualitatif dalam hal tujuan penelitiannya adalah untuk memahami bagaimana suatu komunitas atau individu-individu dalam menerima isu tertentu. Dalam hal ini, sangat penting bagi peneliti yang menggunakan metode kualitatif untuk memastikan kualitas dari proses penelitian, sebab peneliti tersebut akan menginterpretasi data yang telah dikumpulkannya.

Metode kualitatif membantu ketersediaan diskripsi yang kaya atas fenomena. Kualitatif mendorong pemahaman atas substansi dari suatu peristiwa. Dengan demikian, penelitian kualitatif tidak hanya untuk memenuhi keinginan peneliti untuk mendapatkan gambaran/penjelasan, tetapi juga membantu untuk mendapatkan penjelasan yang lebih dalam (Sofaer, 2018) Dengan demikian, dalam penelitian kualitatif, peneliti perlu membekali dirinya dengan pengetahuan yang memadai terkait permasalahan yang akan ditelitinya. Diantaranya sebagai berikut

- A. Peneliti menggunakan prosedur mendapatkan data yang tepat.
- B. Peneliti membatasi penelitian di dalam asumsi dan karakteristik dari pendekatan kualitatif.
- C. Peneliti menggunakan pendekatan kualitatif dalam penelitiannya.
- D. Peneliti memulai penelitian dengan satu fokus.
- E. Penelitian berisi metode yang rinci, pendekatan yang tepat dalam pengumpulan data, analisis data, dan penulisan laporan.
- F. Peneliti menganalisis data menggunakan pemisahan analisis dalam beberapa level.
- G. Peneliti menulis secara persuasif, sehingga pembaca dapat merasakan pengalaman yang sama.
- H. Proses penelitian dengan pendekatan kualitatif

Dalam hal pengumpulan data, (Gill, 2018). mengemukakan terdapat beberapa macam metode pengumpulan data dalam penelitian kualitatif yaitu observasi, analisis visual, studi pustaka, dan interview (individual atau grup). Namun demikian, yang paling populer adalah menggunakan metode interview dan focus group discussion (FGD). Selanjutnya data yang berhasil

dikumpulkan, dianalisis untuk dapat memahami dan mendapatkan kesimpulan dalam penelitian tersebut.

- Pengumpulan Data

Data yang digunakan, yang terdiri dari kebakaran hutan Aljazair, merupakan dataset publik. Data sebaran kebakaran hutan di wilayah Aljazair masuk dalam koleksi. Di situs web bernama UCI Repository & Kaggle untuk pengambilan sampel, Anda bisa mendapatkan data ini (Novianti, 2019).

- Analisis Data

Sistem ini membutuhkan Rapidminer untuk berfungsi. Situs web Repositori UCI disajikan sebagai sumber langsung untuk data yang digunakan dalam penyelidikan ini. Tujuan dari penelitian ini adalah mengkategorikan data dari website UCI Repository dan membandingkan akurasi klasifikasi dari algoritma Naive Bayes, Decision Tree, dan KNN menggunakan software Rapidminer (Mardalius, 2018).

- Tahapan Penelitian

Prosedur penelitian saat ini meliputi langkah-langkah pemilihan data, preprocessing, transformasi data, data mining, dan evaluasi. Prosedur tersebut dapat dijelaskan sebagai berikut:

- a. Persiapan dataset

Buat kumpulan data Langkah selanjutnya adalah menghapus data karena dapat berisi informasi yang telah dihapus atau tidak diperlukan lagi. Tahap ini melibatkan mengubah beberapa data integer menjadi data binomial dan melakukan penambahan data pada data kosong. Tipe data diubah menjadi binomial karena ada dua tipe data yang berbeda, dan label kemudian dibuat di atas meja (Mustikasari, 2019).

- b. Validasi dataset

Setelah pengolahan data, akan dilakukan validasi data untuk mengubah data menjadi informasi untuk tugas dan pelajaran. Cross X Validation akan digunakan untuk memastikan kinerja terbaik dari algoritma yang akan digunakan.

- c. Data mining

Pada penelitian ini yaitu menerapkan model algoritma Naïve bayes , Decision tree dan knn untuk mengklasifikasikan dan menentukan tingkat akurasi algoritma yang lebih baik.

d. Algoritma

Algoritma C45: Pohon keputusan adalah algoritme pembelajaran mesin yang menggunakan proses sederhana untuk membuat kesimpulan dengan struktur yang sangat mirip dengan pohon yang memodelkan kemungkinan keberhasilan, biaya tenaga kerja, potensi manfaat atau risiko, dan variabel lainnya. Strateginya adalah menggunakan algoritma dengan pernyataan yang ditulis dalam teks biasa yang menggunakan komplotan rahasia yang belum sempurna untuk menjelaskan berbagai masalah bernuansa yang dapat diatasi untuk menghasilkan hasil yang bermanfaat (Davita, 2022).

Algoritma Naive Bayes adalah satu-satunya metode klasifikasi statistik yang sekarang digunakan, dan ini menunjukkan kemungkinan bahwa kumpulan data yang relevan akan masuk ke dalam kategori yang relevan. Teori probabilitas adalah. Thomas Bayes memperkenalkan teori Naive Bayes pada 18 Agustus tahun itu. Dalam studi perbandingan, satu-satunya algoritma klasifikasi yang digunakan adalah Bayesian, juga dikenal sebagai Pengklasifikasi Naive Bayes. Klasifikasi menggunakan algoritma Naive Bayes dengan akurasi dan kecepatan tinggi saat menggunakan database yang besar. Metode ini sering digunakan untuk menyelesaikan masalah dengan alat peraga karena kepekaannya yang kuat dan rentang pengaruh yang halus (Pribadi 2018) KNN: Algoritma Nearest Neighbor Retrieval, sering dikenal dengan K-NN atau K-NN, adalah alat untuk mengklasifikasikan objek dengan pengetahuan yang paling dekat hubungannya dengan objek yang sedang diklasifikasikan. Situasi dimana klasifikasi ditentukan berdasarkan data pendidikan terakhir (dekat) (Sutoyo 2018)

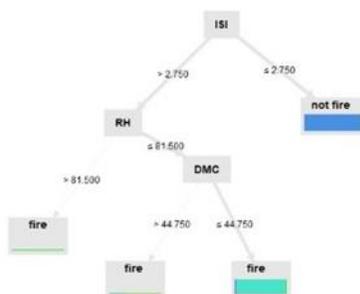
e. Korelasi

Korelasi adalah alat yang digunakan untuk membuka komunikasi antar variabel yang berbeda. Untuk memahami bahwa tidak ada saling ketergantungan yang benar antara variabel, analisis korelatif dilakukan. Jika korelasi variabel ditemukan negatif, hubungan antara variabel dikatakan miring (Mardikaningsih 2021). Namun, jika korelasi antar variabel dipastikan positif dan hub hub antar variabel ditetapkan, maka menunjukkan bahwa data kebakaran hutan Aljazair akurat.

PEMBAHASAN

Data dari kebakaran hutan Aljazair yang tersedia di website UCI digunakan dalam investigasi ini. Kumpulan data yang mencakup kebakaran hutan tahun 2012 di dua lokasi di Aljazair, awalnya disediakan pada tahun 2012 dan akan terus berkembang hingga tahun 2022. Khususnya di wilayah Bejaia dan Sidi Bel-abbes. Peralatan penambangan berkecepatan tinggi digunakan untuk memeriksa data, dan tiga teknik pemrosesan berbeda digunakan untuk melihat mana yang memberikan hasil paling akurat (Ananthi 2022).

Metode yang digunakan untuk penambang cepat menggunakan metode klasifikasi yang menggunakan tiga di antaranya. Decision Tree, k-NN dan Naive Bayes. Ketiga metode ini memiliki tingkat akurasi yang berbeda-beda. Menghasilkan data sebagai berikut:



Gambar 1. Pengujian pertama menggunakan metode klasifikasi Decision Tree.

Berdasarkan temuan pengujian pendekatan pohon keputusan di atas, tidak akan terjadi kebakaran hutan jika indeks sebaran api (ISI) kurang dari atau sama dengan 2.750, namun jika ISI lebih dari 2.750, diperlukan tindakan. Dengan mengamati indikator kelembapan relatif (RH), Anda dapat melakukan kontrol ekstra.

Jika pengukuran indikator RH lebih besar dari 81.500, hal ini mengindikasikan adanya kebakaran hutan; namun, jika kurang dari atau sama dengan 81.500, ini menunjukkan bahwa indikator DMC, atau kelembaban permukaan tanah rata-rata, harus diperiksa.

Dapat disimpulkan dari indikator DMC yang ditampilkan di pohon keputusan bahwa angka indikator akan terus menghasilkan hasil yang sama, yaitu. Kebakaran pohon masih akan terjadi jika nilai indikator DMC lebih besar dari 44.750, kurang dari 44.750, atau tepat 44.750.

Menurut data tersebut, dapat disimpulkan bahwa tingkat penyebaran api yang tidak mulai adalah sekitar 2.750. Kelembaban dan jumlah rata-rata tanah organik yang tidak mudah terbakar dengan kedalaman pembakaran sedang berdampak pada laju penyebaran ini. 42.750 tercatat. Ini

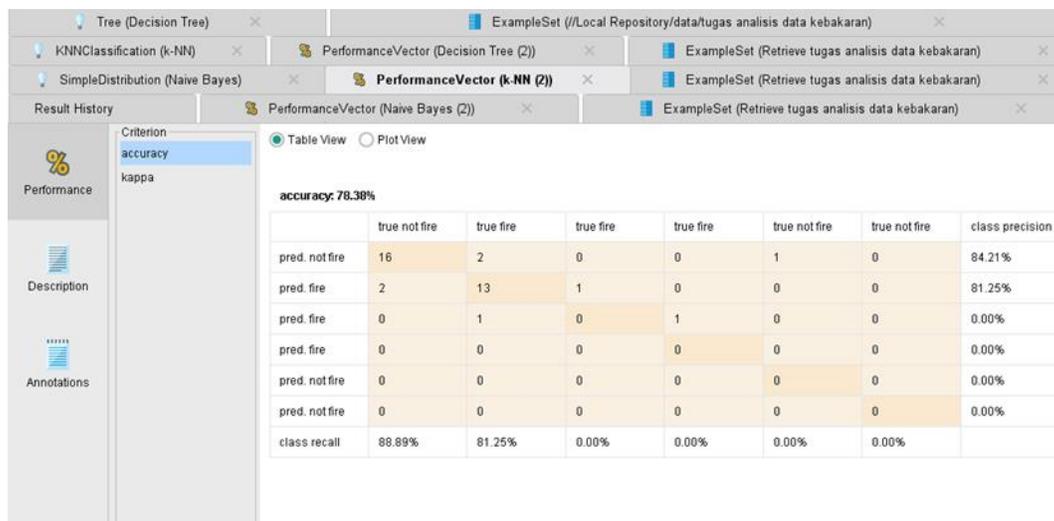
akan menghasilkan 81.500 rekaman kebakaran hutan jika kelembaban adalah satu-satunya faktor yang mempengaruhinya (Fitriyani, 2022).



Gambar 2. Pada metode . Decision tree mendapatkan hasil accuracy sebesar 86.49%

Data yang dapat dilihat dengan metode klasifikasi K-NN hanya memungkinkan untuk klasifikasi berdasarkan jumlah data yang terperkolasi dan pelabelan "terbakar" dan "tidak terbakar", yang mempersulit pembaca untuk mengakses data yang tersedia.

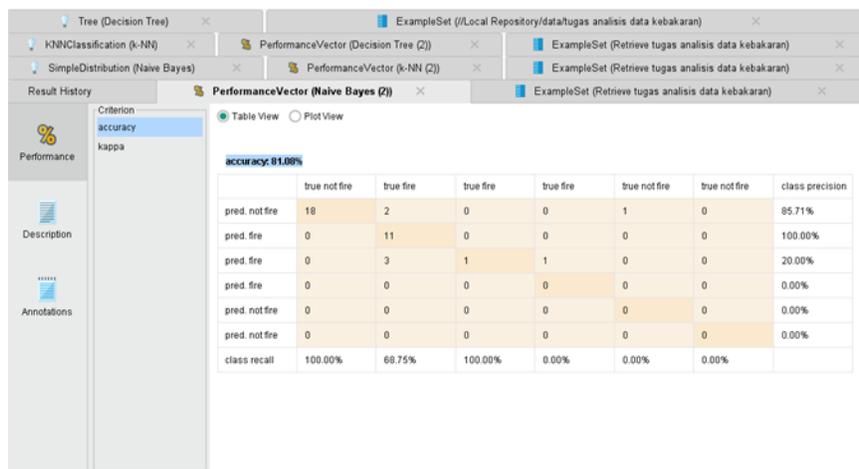
Pada metode ini, hasil akurasi jauh lebih buruk dibandingkan dengan Decision Tree, dengan tingkat akurasi metode K-NN sebesar 78,38%. pada tahun 2022 (Fitriyani 2022).



Gambar 3. Pada pengujian ketiga kami menggunakan metode classifikasi KNN

Yang ketiga, data yang diperoleh dengan metode klasifikasi naive bayes sedikit lebih baik diilustrasikan dengan menunjukkan klasifikasi data yang terbakar dan tidak terbakar, meskipun jumlahnya besar, tetapi terlihat lebih baik daripada metode KNN.

Untuk metode ini, hasil akurasi cukup baik dibandingkan nilai akurasi metode Naive Bayes untuk KNN sebesar 81.08% (Fitriyani, 2022).



Gambar 4. Pada pengujian ketiga kami menggunakan metode classifikasi Naïve Bayes

	Metode pengujian data		
	Decision Tree	k-NN	Naive Bayes
persentasi accuracy	86.49%	78.38%	81.08%

Tabel 1. Table hasil penelitian menggunakan ketiga metode classifikasi

Hal ini dapat disimpulkan dari hasil analisis dataset dengan menggunakan metode klasifikasi Decision Tree, k-NN, dan Naive Bayes. Decision Tree mencapai akurasi tertinggi, dengan akurasi hasil sekitar 86,49%, dan memberikan hasil yang sangat detail mengenai informasi yang digunakan dalam metode tersebut. pada tahun 2022 (Fitriyani, 2022).

SARAN

1. Untuk membatasi kerugian akibat bencana alam, kita harus waspada dalam menjaga hutan untuk mengurangi kerusakan akibat kebakaran hutan.
2. Perlu menghentikan perilaku negatif terkait dengan kecerobohan kita terkait dengan penggunaan api di hutan untuk membuka lahan secara liar dan hal-hal lain yang dapat memicu kebakaran hutan.
3. Pada penelitian berikutnya.penggunaan dataset ini dapat berkermbang sangat pesat dengan riset penelitian aspek yang sangat luas dan dapat bermanfaat bagi kehidupan kedepannya bukan hanya untuk penelitian dengan metode klasifikasi saja melainkan menggunakan metode penelitian lainnya untuk menentukan bagaimana potensi akurasi yang akan di dapatkan oleh dataset ini lebih baik lagi.

PENUTUP

Hal ini dapat disimpulkan dari hasil analisis dataset dengan menggunakan metode klasifikasi Decision Tree, k-NN, dan Naive Bayes. Pohon Keputusan mencapai akurasi tertinggi, dengan akurasi hasil sekitar 86,49%, dan menghasilkan hasil yang sangat rinci mengenai informasi yang digunakan dalam metode tersebut.

Data yang dapat dilihat menggunakan metode klasifikasi K-NN hanya mengizinkan klasifikasi berdasarkan jumlah data yang ditampilkan dan label "terbakar" dan "tidak terbakar", yang membatasi akses ke data yang tersedia bagi pembaca. Data yang dapat dilihat dengan metode klasifikasi K-NN hanya memungkinkan untuk klasifikasi berdasarkan jumlah data yang terperkolasi dan pelabelan "terbakar" dan "tidak terbakar", yang mempersulit pembaca untuk mengakses data yang tersedia. Pada pengujian ketiga menggunakan metode classifikasi naïve bayes data yang didapatkan sedikit lebih tergambar dengan menampilkan klasifikasi data yang terbakar serta tidak terbakar. seberapa besar angkanya, cukup terlihat lebih baik di bandingkan menggunakan metode knn.

Dari data yang kami kumpulkan dapat disimpulkan bahwa penyebaran api-tingkat yang tidak menyebabkan terjadinya kebakaran adalah sekitar 2.750 data, dan penyebaran api-tingkat yang menyebabkan terjadinya kebakaran disebabkan oleh kelembaban udara serta tingkat api-tingkat rata-rata lapisan tanah organik yang tidak menyebabkan kebakaran. Jika hanya kelembaban udara yang diteliti, 81.500 fakta kondisi hutan saat ini terungkap.

DAFTAR PUSTAKA

- A. K. Mishra. (2020). Naive Bayes, Decision Tree and KNN classification algorithms for forest fire prediction using remote sensing data. *International Journal of Geomatics and Geosciences*.
- Aggarwal, C. C. (2018). *Data classification: algorithms and applications*. Springer.
- Alpaydin, E. (2018). *Introduction to machine learning*. Cambridge. Cambridge, MA: MIT Press.
- Amara, H. , & B. N. (2018). A comparative study of decision tree, Naive Bayes and KNN algorithms for forest fire classification in Algeria. . *Journal of Ambient* .
- Ananthi, J., Sengottaiyan, N., Anbukaruppusamy, S., Upreti, K., & Dubey, A. K. (2022). Forest fire prediction using IoT and deep learning. *International Journal of Advanced Technology and Engineering Exploration*, 9(87), 246–256.
- Annissa Widya Davita. (2022, March 28). Pahami Metode Decision Tree Sebagai Algoritma Data Science. DQLab.

- Arbi, M. B. , & O. A. (2021). Evaluation of Naive Bayes, Decision Tree, and KNN Algorithms for Forest Fire Prediction in Algeria. *Journal of Applied Sciences*, 21(2), 105-111.
- Curt, T., Aini, A., & Dupire, S. (2020). Fire activity in Mediterranean forests (The Algerian case). *Fire*, 3(4), 58.
- Djahri, A. , & B. N. (2019). Comparison of Naive Bayes, Decision Tree, and k-Nearest Neighbor Algorithms for Forest Fire Classification in Algeria. *International Journal of Engineering & Technology*.
- Domingos, P. (2018). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. Basic Books.
- Fayyad, U. M. , P.-S. G. , & S. P. (2018). *From data mining to knowledge discovery in databases*. AI Magazine,.
- Fitriyani, F. (2022). Implementasi Forward Selection dan Bagging untuk Prediksi Kebakaran Hutan Menggunakan Algoritma Naïve Bayes. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 8(1), 1–8.
- H. E. Khedr. (2019). orithms for Forest Fire Classification using Remote Sensing Data. *International Journal of Remote Sensing and Earth Sciences*.
- Han, J. , & K. M. (2018). *Data mining. Concepts and Techniques*.
- Handayani, F., & Pribadi, F. S. (2018). Implementasi algoritma naive bayes classifier dalam pengklasifikasian teks otomatis pengaduan dan pelaporan masyarakat melalui layanan call center 110. *Jurnal Teknik Elektro*, 7(1), 19–24.
- Ivanedra, K., & Mustikasari, M. (2019). Implementasi Metode Recurrent Neural Network Pada Text Summarization Dengan Teknik Abstraktif. *J. Teknol. Inf. Dan Ilmu Komput*, 6(4), 377.
- Kaur, K. , & C. K. (2021). (2021). A review of k-nearest neighbor algorithm for data classification. Kaur, K., & Chahal, K. (2021). A Review of k-Nearest Neigh *Journal of Ambient Intelligence and Humanized Computing*, 12(5), 7647-7661.
- Li, H. , & Y. X. (2020). (2019). An improved random forest algorithm for customer churn prediction. *Information Sciences*, 526, 198-212.
- Li, X. , L. Q. , L. H. , & Z. Y. (2019). (2019). Naive Bayesian network-based method for diagnosing knee osteoarthritis. . . *Journal of Medical Systems*, 43(3), 113.
- Lin, C.-H., Yang, P.-K., Lin, Y.-C., & Fu, P.-K. (2020). On machine learning models for heart disease diagnosis. *2020 IEEE 2nd Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS)*, 158–161.
- Mardalius, M. (2018). Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 4(2), 123–132.
- Mardikaningsih, R., Sinambela, E. A., Issalillah, F., Munir, M., & Retnowati, E. (2021). Analisis Korelasi Antara Role Stress dan Komitmen Organisasi. *Jurnal Baruna Horizon*, 4(2), 79–83.
- Mitchell, T. M. (2016). *Machine learning*. McGraw-Hill Education.

- Mohamed Lamine Zargayouna. (2018). A Comparative Study of Naive Bayes, Decision Tree, and KNN Algorithms for Forest Fire Detection. Mustapha Belghit.
- Novianti, D. (2019). Implementasi Algoritma Naïve Bayes Pada Data Set Hepatitis Menggunakan Rapid Miner. *Paradig.-J. Komput. Dan Inform*, 21(1), 49–54.
- S. A. Al-Hussaini. (2021). fire Hazard Assessment Using Naive Bayes, Decision Tree, and KNN Algorithms Based on Remote Sensing Data. *Journal of Forest Research and Development*.
- Suresh Bista, R. S. B. G. (2018). “Classification Algorithms in Forest Fire Analysis: A Review.”
-.
- Sutoyo, M. N. (n.d.). Algoritma K-NN.
- Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung. *Jurnal Media Informatika Budidarma*, 4(2), 437–444.
- Wang, Z. , & J. Y. (2019). (2019). Multi-layer decision tree algorithm based on rough set theory for medical diagnosis. *Journal of Medical Systems*, 43(8), 435.
- Zhang, L. , & C. H. (2019). A Naive Bayes-based imbalanced data classification method using a weighted prior probability. *Neurocomputing*, 342, 10-19.
- Zhang, L. , & C. H. (2019). Decision tree-based imbalanced data classification using a modified distance measure. . *Neurocomputing*, 342, 36-44.
- Gill, P. S. (2018). *Methods of data collection in qualitative research*. Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). *Methods of data collection in British Dental Journal*.
- McCusker, K. &. (2018). *Research using qualitative. quantitative or mixed methods and choice based on the research*.
- Sofaer, S. (2018). *Qualitative methods. Health Services* .