

KLASIFIKASI KEBAKARAN HUTAN DAN LAHAN MENGGUNAKAN ALGORITMA NAÏVE BAYES DI KABUPATEN PELALAWAN

Trya Ayu Pratiwi¹, Muhammad Irsyad², Rahmad Kurniawan³
Surya Agustian⁴, Benny Sukma Negara⁵

^{1,2,3,4,5} Program Studi Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau
Jalan. HR. Soebrantas Km. 15 Panam Pekanbaru, Riau

¹11651203413@students.uin-suska.ac.id, ²irsyadtech@uin-suska.ac.id, ³rahmadkurniawan@uin-suska.ac.id,
⁴surya.agustian@uin-suska.ac.id, ⁵bsnegara@uin-suska.ac.id

Abstrak— Kebakaran hutan dan lahan merupakan salah satu masalah lingkungan dalam hal ekonomis dan ekologis yang merugikan. Jumlah hotspot kebakaran hutan di Provinsi Riau telah meningkat secara dramatis menyebabkan kabut asap yang berbahaya. Penelitian ini bertujuan untuk mengklasifikasi kebakaran hutan dan lahan di Kabupaten Pelalawan, Riau. Penelitian ini telah menggunakan 792 data dalam lima tahun (2015-2019). Model klasifikasi kebakaran hutan menggunakan algoritma Naïve Bayes. Atribut yang digunakan untuk klasifikasi terdiri dari suhu, kelembaban, curah hujan, kecepatan angin, dan kelas dari keempat faktor tersebut didapatkan hasil probabilitas tertinggi yaitu suhu dengan nilai 0.978260870 dan nilai akurasi tertinggi adalah dataset tahun 2017 dengan nilai akurasi adalah 81.03% sehingga nilai akurasi tertinggi tersebut dapat diterapkan pada dataset baru yaitu pada tahun 2019 dengan nilai akurasi nya adalah 82%. Dengan demikian, dapat disimpulkan bahwa model klasifikasi menggunakan Naïve Bayes memiliki potensi untuk digunakan secara efektif sehingga dapat membantu pihak berwenang dalam mencegah kebakaran hutan dan lahan.

Kata Kunci— Kebakaran hutan, Lahan, Klasifikasi, Naïve Bayes, Pelalawan.

Abstract— Forest and land fires are one of the environmental problems in terms of economic and ecological disadvantages. The number of forest fire hotspots in Riau Province has increased dramatically causing a dangerous haze. This study aims to classify forest and land fires in Pelalawan Regency, Riau. This study has used 792 data in five years (2015-2019). The forest fire classification model uses the Naïve Bayes algorithm. The attributes used for classification consist of temperature, humidity, rainfall, wind speed, and the class of the four factors, the highest probability result is temperature with a value of 0.978260870 and the highest accuracy value is the 2017 dataset with an accuracy value of 81.03% so that the highest accuracy value this can be applied to the new dataset, namely in 2019 with an accuracy value of 82%. Thus, it can be concluded that the classification model using Naïve Bayes has the potential to be used effectively so that it can assist the authorities in preventing forest and land fires.

Keywords— Forest fires, Land, Classification, Naïve Bayes, Pelalawan.

I. PENDAHULUAN

Kebakaran hutan dan lahan yang terjadi di Indonesia telah menimbulkan banyak kerugian bagi masyarakat. Luas kebakaran hutan dan lahan di Indonesia telah meningkat secara drastis pada tahun 2015 dengan luas 2.611.411,44 Ha [3]. Luas kebakaran hutan dan lahan terburuk di Indonesia terjadi di Provinsi Riau, Jambi, Sumatera Selatan dan Kalimantan Tengah [3]. Provinsi Riau merupakan salah satu wilayah dengan kebakaran hutan dan lahan yang cukup tinggi dan terjadi setiap tahun. Hasil data luas kebakaran hutan dan lahan di Provinsi Riau dari pantauan *Karhutla Monitoring System*, area yang terbakar pada tahun 2015 yakni sekitar 183.808,59 Ha, pada tahun 2016 sekitar 85.219,51 Ha, pada 2017 sekitar 6.866,09 Ha, pada tahun 2018 sekitar 37.236,27 Ha hingga pada tahun 2019 sekitar 76.267,00 Ha [3]. Kebakaran hutan umumnya terjadi pada bulan Agustus dan September, bertepatan dengan musim kemarau pada sebagian besar

wilayah Indonesia [15]. Curah hujan yang sangat rendah pada musim kemarau mempengaruhi terjadinya kebakaran hutan karena menyebabkan hutan dan lahan mudah terbakar [7].

Dikutip dari www.kompas.com (2019) berbagai upaya pemerintah dalam mencegah terjadinya kebakaran hutan di Provinsi Riau yaitu dengan melakukan *water bombing*, menyediakan masker gratis, meliburkan aktivitas sekolah, tetapi hal tersebut bukan solusi yang tepat untuk jangka panjang [9]. Tindakan yang tegas dan nyata seperti penegakan hukum, dan diperlukan sistem alarm kebakaran yang dapat berfungsi sebagai sistem peringatan dini [16]. Dalam pengembangan sistem peringatan dini, diperlukan hasil dari *Data Mining* yang akurat dan dapat digunakan untuk memberikan informasi yang efektif dalam pencegahan kebakaran hutan. Oleh karena itu, penelitian ini bertujuan untuk menggunakan *Data*

Mining dalam mengklasifikasi titik panas yang ada di Provinsi Riau.

Data Mining adalah proses ekstraksi suatu data yang bersifat implisit dan tidak berguna menjadi informasi dari data yang jumlahnya besar [32]. Salah satu bagian dari *Data Mining* adalah metode klasifikasi [30]. Salah satu penerapan klasifikasi pada kebakaran hutan adalah *K-Nearest Neighbor* untuk klasifikasi indeks cuaca kebakaran berdasarkan data *Automatic Weather Station* [25]. Berdasarkan hasil penelitian Reza Noviansyah (2018) diperoleh bahwa algoritma *K-Nearest Neighbor* akurat digunakan untuk menghasilkan klasifikasi tingkat bahaya kebakaran hutan dengan atribut data berupa suhu, kelembaban, curah hujan dan kecepatan angin, yang menghasilkan akurasi persentase keberhasilan 80,16% dengan nilai $K=5$. Namun algoritma *K-Nearest Neighbor* ini mempunyai kelemahan seperti menentukan nilai K optimal, tidak menangani *missing value* secara implisit, sensitif terhadap *data outlier* dan nilai komputasi yang tinggi [5].

Penelitian lainnya dilakukan oleh Negara & Kurniawan (2019) menggunakan *Decision Tree* dan *Bayesian Network*. Algoritma ini memiliki kelebihan dalam menjelaskan hubungan setiap atribut data cuaca yang mempengaruhi tingkat hotspot menggunakan *Decision Tree* dan *Bayesian Network* memiliki nilai akurasi yang baik dan tingkat kesalahan yang rendah. Namun, algoritma *Decision Tree* memiliki kelemahan seperti terjadi tumpang tindih jika kelas dan kriteria yang digunakan sangat banyak, ini menyebabkan waktu keputusan sangat lama dan membutuhkan banyak memori [2]. Walaupun *Bayesian Network* memiliki kelebihan dan banyak digunakan oleh para peneliti, *Bayesian Network* juga memiliki kelemahan yaitu sulit untuk mencapai kesepakatan beberapa pakar dalam membangun struktur *Bayesian Network*, pakar juga akan memerlukan waktu yang lama hanya untuk membangun struktur *Bayesian Network* dan Pakar akan sulit menentukan *conditional probability* berdasarkan pendapat pakar [19].

Penelitian yang dilakukan oleh Widiastuti (2012) yang membahas tentang perbandingan algoritma akan dilihat berdasarkan nilai *correctly classified instance*, *incorrectly classified*, *kappa statistic*, *true positive*, *false positive*, dan *confusion matrix*. Dalam penelitian ini, algoritma yang dibandingkan adalah algoritma SVM, *Decision Tree*, dan *Naive Bayes*. Dapat disimpulkan algoritma yang memiliki kinerja yang lebih unggul adalah *Decision Tree*.

Berdasarkan tinjauan pustaka, telah dilakukan penelitian terhadap analisis perbandingan algoritma SVM, *Naive Bayes*, dan *Decision Tree* dalam klasifikasi [31]. *Decision Tree* menjadi algoritma terbaik tetapi tidak maksimal dalam waktu komputasi (*running time*), untuk membangun sebuah model algoritma yang tercepat dari ketiga algoritma tersebut adalah *Naive Bayes* [20]. Maka penelitian ini menggunakan algoritma *Naive Bayes* untuk

menganalisis metode klasifikasi untuk memprediksi kemunculan titik panas di Provinsi Riau. Metode klasifikasi *Naive Bayes* dipilih karena memiliki pemodelan dan hasil akurasi yang baik dalam set pelatihan data dan algoritma *Naive Bayes* merupakan algoritma yang paling baik dalam hal waktu komputasi (waktu yang dibutuhkan untuk membangun sebuah model), dan sering bekerja dengan *running time* yang tercepat dengan tingkat akurasi yang tinggi [20].

Penelitian kebakaran hutan dilakukan oleh Misfaul (2018) menggunakan algoritma *Naive Bayes* untuk merancang sistem deteksi titik kebakaran dengan metode *Naive Bayes* menggunakan sensor suhu dan sensor api berbasis *Arduino* dan menghasilkan tiga kesimpulan yaitu pertama sistem pendeteksi lokasi titik kebakaran ketika diuji dalam ruangan tidak menggunakan AC dengan data pengujian sebanyak 36 titik diperoleh keakuratan sebanyak 94%. Kesimpulan kedua ketika sistem pendeteksi kebakaran digunakan pada ruang yang menggunakan AC dari 36 titik pengujian didapatkan keakuratan sebanyak 86%. Kesimpulan ketiga sistem membutuhkan waktu 0.48 detik untuk menentukan keputusan di mana lokasi terjadinya kebakaran [15].

Penelitian selanjutnya dilakukan oleh Rizka (2016) yang membahas tentang klasifikasi kemunculan titik panas pada lahan gambut di Sumatera dan Kalimantan menggunakan algoritma *Naive Bayes* dan menghasilkan dataset dengan nilai akurasi tertinggi untuk wilayah Kalimantan dan Sumatera adalah dataset tahun 2006 dengan nilai akurasi adalah 99.996% untuk Kalimantan dan 99.992% untuk Sumatera. Model dengan akurasi tertinggi berhasil diterapkan pada dataset baru tahun 2015 dengan akurasi 100% untuk Kalimantan dan 99.02% untuk Sumatera. Maka penelitian berhasil mengklasifikasikan kemunculan titik panas di lahan gambut di Sumatera dan Kalimantan menggunakan algoritma *Naive Bayes* [27].

Penelitian selanjutnya dilakukan oleh Nugroho (2017) yang membahas tentang klasifikasi tingkat kerawanan kebakaran hutan dan lahan menggunakan algoritma *Naive Bayes* di Pulang Pisau, Kalimantan Tengah dan berhasil mengklasifikasikan tingkat kerawanan kebakaran hutan dan lahan pada daerah Pulang Pisau, Kalimantan Tengah menggunakan algoritma *Naive Bayes*. Model klasifikasi menghasilkan rata-rata nilai akurasi data latih sebesar 71.02% [19]. Penelitian selanjutnya dilakukan oleh Putrada (2018) yang membahas tentang *Increasing Smoke Classifier Accuracy Using Naive Bayes Method On Internet Of Things* dan menghasilkan proses klasifikasi *Naive Bayes* memiliki rentang tingkat akurasi 88% hingga 91% sehingga hasil ini dapat diterima untuk akurasi klasifikasi [25]. Penelitian selanjutnya dilakukan oleh Andi Nugroho (2019) yang membahas tentang *Peatland Forest Fire Prevention Using Wireless Sensor Network Based on Naive Bayes Classifier* dan *Naive Bayes* digunakan untuk memutuskan apakah kebakaran hutan akan terjadi atau tidak [1].

Penelitian ini diharapkan dapat menghasilkan model yang akurat untuk memprediksi kemungkinan munculnya titik panas di masa yang akan datang sehingga pihak yang berwenang dapat melakukan tindakan pencegahan terjadinya kebakaran hutan dan lahan.

II. LANDASAN TEORI

A. Kebakaran Hutan

Kebakaran hutan dan lahan adalah suatu peristiwa kebakaran yang dilakukan oleh perbuatan manusia maupun alami, yang ditandai oleh penjalaran api serta membakar hutan dan lahan yang dilaluinya [13]. Penelitian kebakaran hutan dilakukan oleh Mulyana (2019) yang membahas tentang bencana kabut asap akibat kebakaran hutan dan lahan serta pengaruh terhadap kualitas udara dan menghasilkan Jumlah titik api kebakaran hutan di Provinsi Riau selama bulan Februari - Maret pada tahun 2014 mencapai 2543 titik api dengan perincian 1319 titik api yang terjadi pada bulan Februari dan 1224 titik api yang terjadi pada bulan Maret. Penelitian selanjutnya dilakukan oleh Setiawan (2018) yang membahas tentang kebijakan penanganan kebakaran hutan dan lahan di Indonesia yang menghasilkan upaya pemerintah dalam pemberdayaan masyarakat dalam pengguna lahan agar tidak membakar hutan sembarangan dan menemukan cara baru agar tidak merusak lingkungan.

Telah dilakukan penelitian di bidang Informatika sebelumnya Reza Noviansyah (2018) yang menggunakan metode klasifikasi tentang penerapan *Data Mining* menggunakan metode *K-Nearest Neighbor* untuk klasifikasi indeks cuaca kebakaran berdasarkan data *Automatic Weather Station* yang menghasilkan klasifikasi tingkat bahaya kebakaran hutan dengan atribut data berupa suhu, kelembaban, curah hujan dan kecepatan angin, menghasilkan akurasi persentase keberhasilan 80,16% dengan nilai $K=5$. Namun algoritma *K-Nearest Neighbor* ini mempunyai kelemahan seperti menentukan nilai K optimal, tidak menangani *missing value* secara implisit, sensitif terhadap *data outlier* dan nilai komputasi yang tinggi [5].

Penelitian lainnya dilakukan oleh Negara & Kurniawan (2019) menggunakan *Decision Tree* dan *Bayesian Network*. Algoritma ini memiliki kelebihan dalam menjelaskan hubungan setiap atribut data cuaca yang mempengaruhi tingkat hotspot menggunakan *Decision Tree* dan *Bayesian Network* memiliki nilai akurasi yang baik dan tingkat kesalahan yang rendah. Namun, algoritma *Decision Tree* memiliki kelemahan seperti terjadi tumpang tindih jika kelas dan kriteria yang digunakan sangat banyak, ini menyebabkan waktu keputusan sangat lama dan membutuhkan banyak memori [2]. Walaupun *Bayesian Network* memiliki kelebihan dan banyak digunakan oleh para peneliti, *Bayesian Network* juga memiliki kelemahan yaitu sulit untuk mencapai kesepakatan beberapa pakar dalam

membangun struktur *Bayesian Network*, pakar juga akan memerlukan waktu yang lama hanya untuk membangun struktur *Bayesian Network* dan Pakar akan sulit menentukan *conditional probability* berdasarkan pendapat pakar [19].

Penelitian yang dilakukan oleh Widiastuti (2012) yang membahas tentang perbandingan algoritma akan dilihat berdasarkan nilai *correctly classified instance*, *incorrectly classified*, *kappa statistic*, *true positive*, *false positive*, dan *confusion matrix*. Dalam penelitian ini, algoritma yang dibandingkan adalah algoritma SVM, *Decision Tree*, dan *Naive Bayes*. Dapat disimpulkan algoritma yang memiliki kinerja yang lebih unggul adalah *Decision Tree*.

Berdasarkan tinjauan pustaka, telah dilakukan penelitian terhadap analisis perbandingan algoritma SVM, *Naive Bayes*, dan *Decision Tree* dalam klasifikasi [31]. *Decision Tree* menjadi algoritma terbaik tetapi tidak dapat maksimal dalam waktu komputasi (*running time*), untuk membangun sebuah model algoritma yang tercepat dari ketiga algoritma tersebut adalah *Naive Bayes* [20]. Maka penelitian ini menggunakan algoritma *Naive Bayes* untuk menganalisis metode klasifikasi untuk memprediksi kemunculan titik panas di Provinsi Riau.

B. Titik Panas

Titik panas dapat digunakan untuk mengidentifikasi adanya kebakaran. Titik panas adalah indikator kebakaran hutan yang mendeteksi suatu lokasi yang memiliki suhu relatif lebih tinggi dibandingkan dengan suhu disekitarnya [22]. Menurut Kementerian Lingkungan Hidup dan Kehutanan [8]. jumlah titik panas tahun lalu hingga September tercatat ada sekitar 4.079 titik panas, sehingga total titik panas tahun 2019 sekitar 7.354 titik panas dan terjadi peningkatan sekitar 3.279 titik panas. Sedangkan di Provinsi Riau terdapat jumlah titik panas yaitu sekitar 138 titik api [3].

Suhu titik panas dapat dihasilkan berdasarkan nilai suhu kecerahan (*brightness temperature*) yang ditangkap oleh sensor pada satelit. Satelit akan mengidentifikasi piksel sebagai titik panas apabila piksel memiliki suhu diatas ambang batas. Ambang batas suhu yang diidentifikasi sebagai titik panas yaitu apabila mencapai $>330^{\circ}\text{K}$ [27]. *Brightness temperature* adalah ukuran deskriptif dari pancaran radiasi dalam bentuk suhu yang dipancarkan ke atas pada bagian atmosfer bumi. *Brightness temperature* merupakan fitur dasar pada citra penginderaan jarak jauh yang di deteksi dari lokasi tertentu dan diukur dalam satuan ukur kelvin [23].

C. Data Mining

Data Mining adalah sebuah istilah yang digunakan untuk menggali pengetahuan dalam big data. *Data Mining* adalah proses ekstraksi suatu data bersifat implisit dan tidak berguna menjadi informasi dari data yang jumlahnya besar [32]. *Data Mining* berguna untuk mengidentifikasi informasi yang bermanfaat dengan

menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning dalam sebuah dataset besar [12]. Definisi lain mengatakan *Data Mining* adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan pola dari suatu data yang berukuran besar [28]. Tipe data yang digunakan dalam *Data Mining* dibedakan menjadi tiga yaitu tipe data numerik, tipe data kategorikal dan tipe data rentang waktu [30].

Menurut buku *Data Mining: Algoritma dan Implementasi dengan Pemrograman PHP* secara umum *Data Mining* terdapat lima peranan berdasarkan tugasnya yaitu sebagai berikut [30]:

- a. *Estimasi*
Estimasi adalah proses memperkirakan nilai yang belum diketahui secara pasti.
- b. *Prediksi*
Prediksi adalah proses memprediksi nilai terhadap suatu variabel yang berlandaskan nilai dari suatu variabel lainnya.
- c. *Klasifikasi*
Klasifikasi adalah proses untuk menemukan model yang menjelaskan atau membedakan tentang konsep atau kelas datanya, dengan tujuan untuk dapat memperkirakan kelas mana yang labelnya tidak diketahui [4].
- d. *Clustering*
Clustering adalah proses meminimumkan jarak yang ada di dalam *cluster* dan jarak antar sesama *cluster* dimaksimumkan.
- e. *Asosiasi*
Association merupakan proses menemukan hubungan atau kaitan antara berbagai kombinasi *item* dari atribut tertentu.
Tahapan dari *Data Mining* ada 5 yaitu sebagai berikut [17]:
 - a. *Data Selection*
Seleksi data merupakan suatu proses dari *Data Mining* sebelum menggali informasi-informasi dalam *Data Mining*. Seleksi data dilakukan dengan memilih himpunan data, sampel data, dan menciptakan himpunan data target.
 - b. *Pre-processing*
Pre-processing adalah proses membersihkan data yang bersifat mengganggu, duplikat, data yang inkonsisten dan *typografi* dalam proses *Data Mining*.
 - c. *Transformation*
Transformation adalah proses melakukan perubahan data menjadi bentuk yang layak untuk proses *mining*.
 - d. *Data Mining*
Data Mining adalah sebuah istilah yang digunakan untuk menggali pengetahuan dalam big data. *Data Mining* berguna untuk mengidentifikasi informasi yang bermanfaat dengan menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning dalam sebuah dataset besar [12].

- e. *Interpretation* atau *Evaluation*

Interpretation atau *evaluation* merupakan pola dari informasi proses *Data Mining* yang bisa di mengerti dan mudah di pahami bagi pengguna, meliputi pemeriksaan pola atau informasi yang ditemukan.

D. *Klasifikasi*

Pada *Data Mining* terdapat beragam metode dalam melakukan klasifikasi, Klasifikasi merupakan suatu kegiatan untuk menemukan suatu model atau fungsi yang dapat membedakan konsep atau kelas data untuk kepentingan tertentu [11]. Klasifikasi adalah proses untuk menemukan model yang menjelaskan atau membedakan tentang konsep atau kelas datanya, dengan tujuan untuk dapat memperkirakan kelas mana yang labelnya tidak diketahui [4]. Tujuan dari algoritma klasifikasi adalah untuk memprediksi kelas baru dari *dataset* yang mempunyai kelas [30].

Metode klasifikasi yang terdapat pada *Data Mining* antara lain [6]:

1. *Neural Network*

Neural Network (Jaringan Syaraf Tiruan) adalah prosesor tersebar paralel yang sangat besar dan memiliki kecenderungan untuk menyimpan pengetahuan yang bersifat pengalaman dan membuatnya siap untuk digunakan [6].

2. *Decision Tree*

Decision Tree sendiri merupakan metode klasifikasi dan prediksi yang sangat kuat dan banyak di minati [33].

3. *Naïve Bayes*

Algoritma *Naïve Bayes* adalah salah satu algoritma klasifikasi berdasarkan teorema *Bayesian* pada statistika untuk memprediksi probabilitas suatu kelas [30].

4. *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut, ketepatan algoritma K-NN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan, atau jika bobot fitur tersebut tidak setara dengan relevansi nya terhadap klasifikasi.

5. *Logistic Regression*

Logistic regression adalah bagian dari analisis regresi yang digunakan ketika variabel dependen (respon) merupakan variabel dikotomi.

E. *Naïve Bayes*

Algoritma *Naïve Bayes* adalah salah satu algoritma klasifikasi berdasarkan teorema Bayesian pada statistika untuk memprediksi probabilitas suatu kelas [30]. Penelitian kebakaran hutan dilakukan oleh Misfaul (2018) menggunakan algoritma *Naïve Bayes* untuk merancang sistem deteksi titik kebakaran dengan metode *Naïve Bayes* menggunakan sensor suhu dan sensor api berbasis *Arduino* dan menghasilkan tiga kesimpulan yaitu pertama sistem pendeteksi lokasi titik

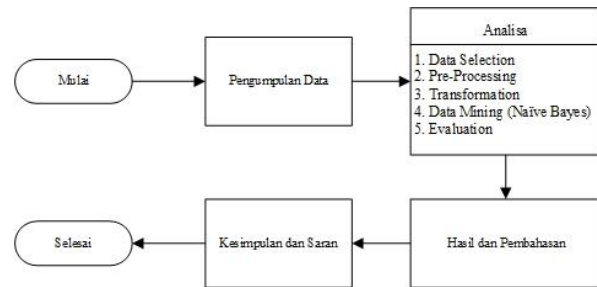
kebakaran ketika diuji dalam ruangan tidak menggunakan AC dengan data pengujian sebanyak 36 titik diperoleh keakuratan sebanyak 94%. Kesimpulan kedua ketika sistem pendeteksi kebakaran digunakan pada ruang yang menggunakan AC dari 36 titik pengujian didapatkan keakuratan sebanyak 86%. Kesimpulan ketiga sistem membutuhkan waktu 0.48 detik untuk menentukan keputusan di mana lokasi terjadinya kebakaran [15].

Penelitian selanjutnya dilakukan oleh Rizka (2016) yang membahas tentang klasifikasi kemunculan titik panas pada lahan gambut di Sumatera dan Kalimantan menggunakan algoritma *Naïve Bayes* dan menghasilkan dataset dengan nilai akurasi tertinggi untuk wilayah Kalimantan dan Sumatera adalah dataset tahun 2006 dengan nilai akurasi adalah 99.996% untuk Kalimantan dan 99.992% untuk Sumatera. Model dengan akurasi tertinggi berhasil diterapkan pada dataset baru tahun 2015 dengan akurasi 100% untuk Kalimantan dan 99.02% untuk Sumatera. Maka penelitian berhasil mengklasifikasikan kemunculan titik panas di lahan gambut di Sumatera dan Kalimantan menggunakan algoritma *Naïve Bayes* [27].

Penelitian selanjutnya dilakukan oleh Nugroho (2017) yang membahas tentang klasifikasi tingkat kerawanan kebakaran hutan dan lahan menggunakan algoritma *Naïve Bayes* di Pulang Pisau, Kalimantan Tengah dan berhasil mengklasifikasikan tingkat kerawanan kebakaran hutan dan lahan pada daerah Pulang Pisau, Kalimantan Tengah menggunakan algoritma *Naïve Bayes*. Model klasifikasi menghasilkan rata-rata nilai akurasi data latih sebesar 71.02%. Penelitian selanjutnya dilakukan oleh Putrada (2018) yang membahas tentang *Increasing Smoke Classifier Accuracy Using Naïve Bayes Method On Internet Of Things* dan menghasilkan proses klasifikasi *Naïve Bayes* memiliki rentang tingkat akurasi 88% hingga 91% sehingga hasil ini dapat diterima untuk akurasi klasifikasi [19]. Penelitian selanjutnya dilakukan oleh Andi Nugroho (2019) yang membahas tentang *Peatland Forest Fire Prevention Using Wireless Sensor Network Based on Naïve Bayes Classifier* dan *Naïve Bayes* digunakan untuk memutuskan apakah kebakaran hutan akan terjadi atau tidak [1]. Maka Sistem Peringatan Dini ini membantu pihak berwenang untuk memantau dan mendeteksi potensi kebakaran hutan di lahan gambut, sehingga dapat dicegah.

III. METODOLOGI PENELITIAN

Metodologi penelitian adalah tahapan atau proses yang disusun secara sistematis dan logis dalam melakukan suatu penelitian yang berguna untuk mencapai target yang diharapkan. Tahapan penelitian yang akan dilakukan pada penelitian ini yaitu sebagai berikut:



Gbr. 1 Tahapan penelitian

A. Perumusan Masalah

Perumusan masalah adalah tahap utama dari metodologi penelitian. Dalam penelitian yang akan dilakukan rumusan masalahnya yaitu bagaimana menganalisis metode klasifikasi untuk memprediksi kemunculan titik panas menggunakan algoritma *Naïve Bayes* sehingga diharapkan dapat menghasilkan model yang akurat.

B. Pengumpulan Data

Pengumpulan data dilakukan dengan dua tahap yaitu observasi dan studi pustaka. Penjelasan dari pengumpulan data adalah sebagai berikut:

1. Observasi

Pada tahapan observasi ini, dilakukan dengan mengumpulkan data kebakaran hutan. Pengumpulan data dapat dilakukan dengan data primer dan data sekunder. Berikut ini penjelasan dari data primer dan data sekunder yaitu:

- a. Data primer merupakan data Kabupaten Pelalawan pada tahun 2015 sampai 2019 yang diambil dari Badan Meteorologi Klimatologi Dan Geofisika Provinsi Riau dengan format .xlsx. Data tersebut digunakan sebagai data training dan data testing. Atribut yang terdapat pada penelitian ini yaitu suhu, curah hujan, kecepatan angin, dan kelembaban. Jumlah dataset kebakaran hutan dan lahan yang digunakan selama 5 tahun adalah 792 data.
- b. Data sekunder yaitu data yang ambil dari Data titik panas (hotspot) untuk seluruh wilayah Indonesia dari citra satelit MODIS Terra dan Aqua yang bersumber dari LAPAN (Lembaga Penerbangan dan Antariksa Nasional) dengan format .csv (LAPAN, 2019).

2. Studi Literatur

Studi pustaka atau biasa disebut dengan *literature* adalah tahapan awal dalam teknik pengumpulan data pada suatu penelitian. Metode ini akan dilakukan dengan cara mencari sumber-sumber referensi terhadap berbagai buku, karya ilmiah maupun jurnal-jurnal yang di anggap penting dan mendukung.

C. Analisa

Analisa adalah tahapan yang dilakukan setelah pengumpulan data dari metodologi penelitian. Analisa merupakan metode khusus yang digunakan untuk menganalisis masalah. Pada tahapan analisa proses

Data Mining peneliti akan menjelaskan tentang bagaimana tahapan-tahapan untuk mencari klasifikasi dari sejumlah data kebakaran hutan dengan menggunakan algoritma *Naïve Bayes* sebagai berikut:

1. *Data Selection*

Pada tahap ini dilakukan tahapan seleksi data. Pemilihan data dalam proses seleksi menggunakan 5 atribut yaitu suhu, curah hujan, kecepatan angin, kelembaban dan kelas.

2. *Pre-processing*

Tahapan untuk pembersihan terhadap data-data yang terdapat data kosong (*missing value*), inkonsisten data, dan ketidaksesuaian data (*outlier*).

3. *Transformation*

Transformasi data dilakukan dalam penelitian ini untuk mengubah tipe data atribut yang memiliki tipe data numerik menjadi tipe data nominal. Hasil dari transformasi tipe data numerik menjadi tipe data nominal ditunjukkan pada Tabel 1 berikut ini:

TABEL 1
TRANSFORMASI DATA

Atribut	Data Numerik	Data Nominal
Suhu	<25°	Rendah
	26°-29°	Sedang
	>29°	Tinggi
Kelembaban	<70-75	Kering
	76-80	Sedang
	80-85	Lembab
	>85	Basah
Curah Hujan	0	Berawan
	0.5-20 mm/hari	Ringan
	20-50 mm/hari	Sedang
	50-100 mm/hari	Lebat
Kecepatan angin (Knot)	1-4	Tenang
	4-7	Sedang
	7-11	Tinggi
	>11	Badai
Kelas	0-30	Tidak kebakaran
	31-80	Kebakaran

4. *Data Mining*

Tahap awal dari proses perhitungan *Naïve Bayes* adalah dengan melakukan pengambilan *data training* dari data kebakaran hutan dan lahan. Adapun variabel penentu yang digunakan dalam mengklasifikasikan data kebakaran hutan dan lahan yaitu:

1. Suhu merupakan variabel yang dikelompokkan dalam tiga kategori yaitu rendah, sedang, dan tinggi.
2. Kelembaban merupakan variabel yang di kelompokkan dalam empat kategori yaitu kering, sedang, lembab dan basah.

3. Curah hujan merupakan variabel yang dikelompokkan dalam empat kategori yaitu berawan, ringan, sedang dan lebat.
4. Kecepatan angin merupakan variabel yang dikelompokkan dalam empat kategori yaitu tenang, sedang, tinggi dan badai.

Berikut merupakan data kebakaran hutan dan lahan pada tahun 2015 dan prediksi potensi kebakaran sebelum dilakukan proses *Data Mining*:

No	Suhu	kelembaban	curah hujan	kecepatan angin	Kelas
1	Sedang	sedang	ringan	tinggi	kebakaran
2	Sedang	lembab	ringan	badai	tidak kebakaran
3	Sedang	lembab	berawan	badai	kebakaran
4	Sedang	lembab	ringan	badai	tidak kebakaran
5	Sedang	sedang	berawan	tinggi	kebakaran
...
149	Sedang	lembab	ringan	tinggi	tidak kebakaran

Gbr. 2 Dataset tahun 2015

Berdasarkan dari tahapan *Data Mining* untuk algoritma *Naïve Bayes*, langkah-langkah dari algoritma *Naïve Bayes* adalah sebagai berikut:

a. Hitung jumlah $class/label = P(C_i)$

$$P(\text{Class kelas} = \text{"tidak kebakaran"}) = 107/149 = 0,718120805$$

$$P(\text{Class kelas} = \text{"kebakaran"}) = 42/149 = 0,281879195$$

b. Menghitung probabilitas $(P(X|C_i))$ kemunculan setiap nilai untuk tiap atribut

Suhu	tidak kebakaran	kebakaran	P(tidak kebakaran)	P(kebakaran)
Rendah	3	1	0,028037383	0,023809524
Sedang	102	39	0,953271028	0,928571429
Tinggi	2	2	0,018691589	0,047619048
Total	107	42		

Kelembaban	tidak kebakaran	Kebakaran	P(tidak kebakaran)	P(kebakaran)
Kering	10	4	0,093457944	0,095238095
Sedang	35	17	0,327102804	0,404761905
Lembab	53	19	0,495327103	0,452380952
Basah	9	2	0,08411215	0,047619048
Total	107	42		

Curah hujan	tidak kebakaran	Kebakaran	P(tidak kebakaran)	P(kebakaran)
Berawan	65	28	0,607476636	0,666666667
Ringan	38	13	0,355140187	0,30952381
Sedang	4	0	0,045045045	0,02173913
Lebat	0	1	0,009009009	0,043478261
Total	107	42		

Pada model *Naive Bayes* tersebut terdapat nilai peluang 0, maka untuk menangani kasus ini dapat digunakan salah satu teknik *Laplace estimation* yaitu pada setiap perhitungan datanya ditambah 1 (satu) dan tidak akan membuat perbedaan yang berarti pada estimasi probabilitas sehingga bisa menghindari kasus nilai probabilitas 0.

Berikut implementasi *Laplace estimation* untuk atribut curah hujan:

$$P(\text{Curah hujan} = \text{"sedang"} \mid \text{kelas} = \text{"tidak kebakaran"}) = \frac{4+1}{107+4} = 0,045045045$$

$$P(\text{Curah hujan} = \text{"sedang"} \mid \text{kelas} = \text{"kebakaran"}) = \frac{0+1}{42+4} = 0,02173913$$

$$P(\text{Curah hujan} = \text{"lebat"} \mid \text{kelas} = \text{"tidak kebakaran"}) = \frac{0+1}{107+4} = 0,009009009$$

$$P(\text{Curah hujan} = \text{"lebat"} \mid \text{kelas} = \text{"kebakaran"}) = \frac{1+1}{42+4} = 0,043478261$$

kecepatan angin	tidak kebakaran	kebakaran	P(tidak kebakaran)	P(kebakaran)
Tenang	3	2	0,028037383	0,047619048
Sedang	8	3	0,074766355	0,071428571
Tinggi	51	22	0,476635514	0,523809524
Badai	45	15	0,420560748	0,357142857
Total	107	42		

c. Kalikan $P(X|Ci) \cdot P(Ci)$

$$P(\text{suhu} = \text{"tinggi"} \mid \text{kelas} = \text{"tidak kebakaran"}) \cdot P(\text{kelembaban} = \text{"kering"} \mid \text{kelas} = \text{"tidak kebakaran"}) \cdot P(\text{curah hujan} = \text{"ringan"} \mid \text{kelas} = \text{"tidak kebakaran"}) \cdot P(\text{kecepatan angin} = \text{"tinggi"} \mid \text{kelas} = \text{"tidak kebakaran"}) \cdot P(\text{kebakaran}) = 0,018691589 \cdot 0,093457944 \cdot 0,355140187 \cdot 0,476635514 \cdot 0,71812080 = 0,000212347$$

$$P(\text{suhu} = \text{"tinggi"} \mid \text{kelas} = \text{"kebakaran"}) \cdot P(\text{kelembaban} = \text{"kering"} \mid \text{kelas} = \text{"kebakaran"}) \cdot P(\text{curah hujan} = \text{"ringan"} \mid \text{kelas} = \text{"kebakaran"}) \cdot P(\text{kecepatan angin} = \text{"tinggi"} \mid \text{kelas} = \text{"kebakaran"}) \cdot P(\text{kebakaran}) = 0,047619048 \cdot 0,095238095 \cdot 0,30952381 \cdot 0,523809524 \cdot 0,281879195 = 0,000207263$$

d. Bandingkan Hasil Per Kelas

Kelas dengan nilai $P(X|Ci) \cdot P(Ci)$ tertinggi adalah kelas yang terpilih, sehingga dapat disimpulkan bahwa status tersebut termasuk dalam klasifikasi "Tidak kebakaran".

5. Evaluation

Penelitian ini menggunakan software R Studio Version 1.2.5042 dalam pencarian pola atau informasi baru yang menarik dan bermanfaat.

IV. HASIL DAN PEMBAHASAN

Pada tahap pembahasan penulis menggunakan pengujian dengan *tools R Studio* untuk melakukan pengecekan *confusion matrix* terhadap klasifikasi yang telah dilakukan oleh algoritma *Naïve Bayes*.

A. Pembagian Data

Proses pembagian data *training* dan data *testing* yang dilakukan di *R Studio* dengan menggunakan *split validation* menjadi data testing dan data training untuk dilakukan analisis. Pada data yang digunakan akan dibagi menjadi 25 % data testing dan 75% data training.

B. Pembuatan Model Klasifikasi

Dalam pembuatan metode klasifikasi menggunakan beberapa *package* yaitu *e1071* dan *caret*. Metode yang telah dibangun dapat memprediksi kelas dan dapat melihat nilai akurasi terhadap metode tersebut. Berikut adalah contoh pernyataan R untuk

membangun model klasifikasi kebakaran hutan dan lahan menggunakan algoritma *Naïve Bayes*:

```
[1] library(readxl)
[2] data = read_excel(file.choose())
[3] View(data)
[4] library(e1071)
[5] library(caret)
[6] str(data)
[7] sampel = sample(1:nrow(data), 0.75*nrow(data), replace = TRUE)
[8] train = data.frame(data)[sampel,]
[9] test = data.frame(data)[-sampel,]
[10] model = naiveBayes(kelas ~ ., data = train, laplace = 1)
[11] model
[12] prediksi = predict(model, test)
[13] hasil = confusionMatrix(table(prediksi, test$kelas))
[14] hasil
```

Package yang digunakan adalah *package e1071* yang memproses model *Naïve Bayes* sedangkan *package caret* digunakan untuk melihat model yang telah dibuat baik atau tidak terdapat *akurasi, sensitivity, specificity*. Program ini akan memproses model *Naïve Bayes* menggunakan *split validation*. Model *naïve bayes* menggunakan *laplace* untuk menangani nilai probabilitas 0 (nol) sehingga setiap perhitungan datanya ditambah 1 (satu). Berikut hasil model berfungsi untuk mencetak hasil model klasifikasi yang berisi probabilitas dan nilai akurasi:

```
> model
Naive Bayes Classifier for Discrete Predictors
Call:
naiveBayes.default(x = X, y = Y, laplace = laplace)
A-priori probabilities:
Y
  kebakaran tidak kebakaran
0.3603604      0.6396396
Conditional probabilities:
suhu
Y
  kebakaran rendah sedang tinggi
tidak kebakaran 0.04225352 0.94366197 0.05633803
kelembaban
Y
  kebakaran basah kering lembab sedang
tidak kebakaran 0.05633803 0.07042254 0.50704225 0.42253521
curah.hujan
Y
  kebakaran berawan ringan
tidak kebakaran 0.6338028 0.3943662
kecepatan.angin
Y
  kebakaran badai sedang tenang tinggi
tidak kebakaran 0.38028169 0.11267606 0.02816901 0.53521127
```

```
> hasil
Confusion Matrix and Statistics

prediksi FALSE TRUE
FALSE      39    14
TRUE       12     2

Accuracy : 0.6119
95% CI : (0.485, 0.7286)
No Information Rate : 0.7612
P-Value [Acc > NIR] : 0.9979

Kappa : -0.1152

McNemar's Test P-value : 0.8445

Sensitivity : 0.7647
Specificity : 0.1250
Pos Pred Value : 0.7358
Neg Pred Value : 0.1429
Prevalence : 0.7612
Detection Rate : 0.5821
Detection Prevalence : 0.7910
Balanced Accuracy : 0.4449

'Positive' class : FALSE
```

C. Perhitungan Akurasi

Perhitungan akurasi dapat dilakukan dengan menggunakan *confusion matriks*:

TABEL 2
CONFUSION MATRIX UNTUK KLASIFIKASI DUA KELAS

Actual	Predicted	
	Kebakaran	Tidak kebakaran
Kebakaran	TN	FP
Tidak kebakaran	FN	TP

Secara umum, evaluasi dapat dirumuskan sebagai berikut :

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Keterangan :

TN = nilai *true negatives*

TP = nilai *true positives*

FP = nilai *false positives*

FN = nilai *false negatives*

TN adalah jumlah kebakaran yang berhasil diprediksi sebagai kebakaran. TP adalah jumlah tidak kebakaran yang berhasil diprediksi tidak kebakaran. FP adalah jumlah kebakaran yang diprediksi tidak kebakaran. Sedangkan, FN menyatakan tidak kebakaran yang diprediksi kebakaran. Akurasi terbaik untuk wilayah kabupaten Pelalawan adalah pada tahun 2015 dengan nilai akurasi 99.45% dan kappa 98.82% dapat dilihat pada Tabel 3 berikut:

TABEL 3
AKURASI MODEL PADA DATASET

Dataset	TP	TN	FP	FN	Akurasi
2015	136	228	2	0	61.19%
2016	132	232	2	1	78.87%
2017	107	257	1	1	81.03%
2018	147	216	2	1	64.02%

D. Penerapan Model dengan Data Baru

Pada tahapan ini akan dilakukan validasi model terhadap dataset dari tahun 2015-2018. Kemudian diterapkan dengan model akurasi tertinggi pada dataset baru tahun 2019. Model dengan akurasi tertinggi terdapat pada tahun 2017. Oleh karena itu, data yang digunakan sebagai data latih merupakan dataset pada tahun 2017, sedangkan data yang digunakan sebagai data uji adalah dataset pada tahun 2019. Berikut sintaks untuk menerapkan model pada data baru tahun 2019:

```
[1] library(readxl)
[2] data = read_excel(file.choose())
[3] View(data)
[4] library(e1071)
[5] library(caret)
[6] str(data)
[7] sampel = sample(1:nrow(data), 0.75*nrow(data), replace = TRUE)
[8] train = data.frame(data)[sampel,]
[9] test = data.frame(data)[-sampel,]
[10] model = naiveBayes(kelas ~ ., data = train, laplace = 1)
[11] model
[12] prediksi = predict(model, test)
[13] hasil = confusionMatrix(table(prediksi, test$kelas))
[14] hasil
```

Package yang digunakan merupakan *e1071*, *package* ini membangun model menggunakan fungsi *Naive Bayes* yang terdapat pada baris keempat pada program dengan pembagian data dilakukan secara *split validation*. Berikut sintaks hasil nilai akurasi pada dataset baru 2019 yaitu sebesar 82%.

```
> hasil
Confusion Matrix and Statistics

prediksi FALSE TRUE
FALSE      119    16
TRUE        11     4

Accuracy : 0.82
95% CI : (0.749, 0.8779)
No Information Rate : 0.8667
P-Value [Acc > NIR] : 0.9596

Kappa : 0.129

McNemar's Test P-value : 0.4414

Sensitivity : 0.9154
Specificity : 0.2000
Pos Pred Value : 0.8815
Neg Pred Value : 0.2667
Prevalence : 0.8667
Detection Rate : 0.7933
Detection Prevalence : 0.9000
Balanced Accuracy : 0.5577

'Positive' class : FALSE
```

Hasil dari *confusion matriks* di atas menunjukkan bahwa model klasifikasi data baru terdapat 199 data kebakaran(F) yang diklasifikasikan sebagai data kebakaran (F) sedangkan terdapat 16 data tidak kebakaran(T) yang diklasifikasikan sebagai data kebakaran(F) terdapat 11 data tidak kebakaran(T) yang diklasifikasikan sebagai data kebakaran(F) dan terdapat

4 data tidak kebakaran(T) yang diklasifikasikan sebagai data tidak kebakaran(T). Dari hasil ini dapat disimpulkan bahwa kinerja Naive Bayes terhadap kebakaran hutan di Kabupaten Pelalawan sangat baik.

TABEL 8
Matriks Confusion Untuk Model pada Dataset

predictive/actual	False (kebakaran)	True (tidak kebakaran)
False (kebakaran)	199	16
True (tidak kebakaran)	11	4

Algoritma *Naive Bayes* menjelaskan ada empat faktor utama yang menyebabkan kebakaran hutan dan lahan yaitu suhu, curah hujan, kecepatan angin dan kelembaban. Probabilitas setiap atribut terhadap kelas kebakaran, dapat mempengaruhi kebakaran hutan dan lahan di kabupaten Pelalawan saat:

1. Suhu dengan probabilitas sedang 0.978260870 yaitu 26-29°
2. Curah hujan dengan probabilitas berawan 0.80434783 yaitu 0 (di)
3. Kecepatan angin dengan probabilitas 0.54347826 yaitu 7-11 knot
4. Kelembaban dengan probabilitas sedang 0.44565217 yaitu 76-80%

Maka kesimpulan yang paling mempengaruhi kebakaran hutan dan lahan di kabupaten Pelalawan adalah atribut suhu dengan probabilitas paling tinggi diantara faktor lainnya yaitu 0.978260870.

```
> model
Naive Bayes Classifier for Discrete Predictors
call:
naiveBayes.default(x = X, y = Y, laplace = laplace)
A-priori probabilities:
Y
  kebakaran tidak kebakaran
  0.1548822      0.8451178
Conditional probabilities:
Y
  suhu
  kebakaran      rendah      sedang      tinggi
  0.043478261    0.978260870    0.010869565
  tidak kebakaran 0.003984064    0.940239044    0.061752988
Y
  kelembaban
  kebakaran      basah      kering      lembab      sedang
  0.13043478    0.06521739    0.40217391    0.44565217
  tidak kebakaran 0.13745020    0.10557769    0.43027888    0.33466135
Y
  curah.hujan
  kebakaran      berawan      lebat      ringan      sedang
  0.80434783    0.03260870    0.19565217    0.01086957
  tidak kebakaran 0.61135378    0.01394422    0.32470120    0.05776892
Y
  kecepatan.angin
  kebakaran      badai      sedang      tenang      tinggi
  0.33695652    0.06521739    0.09782609    0.54347826
  tidak kebakaran 0.27490040    0.13745020    0.20119522    0.39442231
```

V. PENUTUP

Berdasarkan penelitian yang telah dilakukan maka dapat diambil kesimpulan sebagai berikut:

1. Penelitian ini berhasil mengklasifikasikan kemunculan titik panas kebakaran hutan di Kabupaten Pelalawan pada tahun 2015 hingga 2019 menggunakan algoritma *Naive Bayes*.

2. Atribut yang digunakan untuk klasifikasi terdiri dari suhu, kelembaban, curah hujan, kecepatan angin, dan kelas yang menghasilkan nilai akurasi tertinggi adalah dataset tahun 2017 dengan nilai akurasi adalah 81.03% sehingga nilai akurasi tertinggi tersebut dapat diterapkan pada dataset baru yaitu pada tahun 2019 dengan nilai akurasi adalah 82%.
3. Berdasarkan perhitungan Data Mining menggunakan algoritma *Naive Bayes*, dapat ditarik kesimpulan bahwa ada empat faktor utama yang menyebabkan kebakaran hutan dan lahan yaitu suhu, curah hujan, kecepatan angin dan kelembaban dan yang paling mempengaruhi kebakaran hutan dan lahan di kabupaten Pelalawan adalah atribut suhu dengan probabilitas paling tinggi yaitu 0.978260870.
4. Metode klasifikasi menggunakan algoritma *Naive Bayes* dapat digunakan untuk memprediksi kemunculan titik panas sehingga dapat melakukan tindakan pencegahan sebelum terjadi nya kebakaran hutan dan lahan pada masa yang akan datang.

Penelitian selanjutnya diharapkan dapat membangun suatu aplikasi berbasis web untuk menampilkan hasil klasifikasi titik panas menggunakan algoritma *Naive Bayes*.

REFERENSI

- [1] Andi Nugroho, A., Iwan, I., Iroh Nur Azizah, K., & Hakim Raswa, F. (2019). Peatland Forest Fire Prevention Using Wireless Sensor Network Based on *Naive Bayes Classifier*. *KnE Social Sciences*, 2019, 20–34. <https://doi.org/10.18502/kss.v3i23.5134>
- [2] Asfihan, A. (2019). *Decision Tree Adalah : Jenis, Manfaat, Kelebihan dan Kekurangannya*.
- [3] BNPB. (2019). Rekapitulasi Luas Kebakaran Hutan dan Lahan (Ha) Per Provinsi Di Indonesia Tahun 2011-2018. *SiPongi - Karhutla Monitoring Sistem*, (1), 26–27. <https://doi.org/10.1103/PhysRevB.65.020504>
- [4] Bustami. (2014). *PENERAPAN ALGORITMA NAIVE BAYES*. 8(1).
- [5] Cahya. (2018). *Implementasi Data Mining Algoritma k-Nearest Neighbors (k-NN) menggunakan PHP dan MySQL untuk memprediksi kelulusan mahasiswa tepat waktu*.
- [6] Dewi, S. (2016). Komparasi 5 Metode Algoritma Klasifikasi Data Mining Pada Prediksi Keberhasilan Pemasaran Produk Layanan Perbankan. *None*, 13(1), 60–66.
- [7] Itsnaini, N., Sasmito, B., Sukmono, A., & Prasasti, I. (2017). Analisis Hubungan Curah Hujan Dan Parameter Sistem Peringkat Bahaya Kebakaran (Spbk) Dengan Kejadian Kebakaran Hutan Dan Lahan Untuk Menentukan Nilai Ambang Batas Kebakaran. *Jurnal Geodesi Undip*, 6(2), 62–70.
- [8] KLHK. (2019). *Kementerian Lingkungan Hidup dan Kehutanan*.
- [9] Kompas.com. (2019). 4 Upaya Pemerintah Padamkan Kebakaran Hutan. Retrieved from <https://www.kompas.com/tren/read/2019/09/17/213000265/4-upaya-pemerintah-padamkan-kebakaran-hutan?page=all#page3>. website:
- [10] <https://www.kompas.com/tren/read/2019/09/17/213000265/4-upaya-pemerintah-padamkan-kebakaran-hutan?page=all#page3>.
- [11] LAPAN. (2019). Lembaga Penerbangan dan Antariksa Nasional. 2016. Informasi Titik Panas (Hotspot) Kebakaran

- Hutan/Lahan. Retrieved from <https://modis-catalog.lapan.go.id/monitoring/>
- [12] Laroussi, H. M. M. A. (2015). *Implementasi algoritma naïve bayes sebagai proses seleksi penerima beasiswa libyan embassy berbasis web*.
- [13] Maulida, L. (2018). Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov . Dki Jakarta Dengan K-Means. *JISKa*, 2(3), 167–174.
- [14] Menhut. (2016). *Peraturan Menteri Kehutanan Republik Indonesia Nomor P.32 Tahun 2016 tentang Pengendalian Kebakaran Hutan*. Jakarta (ID): Menteri Kehutanan Republik Indonesia.
- [15] Misfaul, M., Dana, M., Kurniawan, W., & Fitriyah, H. (2018). *Rancang Bangun Sistem Deteksi Titik Kebakaran Dengan Metode Naive Bayes Menggunakan Sensor Suhu dan Sensor Api Berbasis Arduino*. 2(9), 3384–3390.
- [16] Mulyana, E. (2019). Bencana Kabut Asap Akibat Kebakaran Hutan Dan Lahan Serta Pengaruhnya Terhadap Kualitas Udara Di Provinsi Riau Februari – Maret 2014. *Jurnal Sains Dan Teknologi Indonesia*, 16(3), 1–7. <https://doi.org/10.29122/jsti.v16i3.3417>
- [17] Negara, B. S., & Kurniawan, R. (2019). *Riau Forest Fire Prediction using Supervised Machine Learning*. (September 2019).
- [18] Novriansyah, D. (2015). *Konsep Data Mining VS Sistem Pendukung Keputusan*. Yogyakarta: Deepublish.
- [19] Nugroho, P. (2017). *Klasifikasi tingkat kerawanan kebakaran hutan dan lahan menggunakan algoritme naïve bayes di pulau pisau, kalimantan tengah*.
- [20] Oktaviana, W. (2017). "Optimasi basis pengetahuan menggunakan algoritma FP-Growth untuk membangun struktur Bayesian Network Penyakit mata. 1–5.
- [21] Pattekari, S.A.; Parveen, A. (2012). Prediction system for heart disease using Naïve Bayes. *International Journal of Advanced Computer and Mathematical Sciences*, 3(3), 290–294.
- [22] Pei J, Kamber M, H. J. (2012). *Data Mining: Concepts and Techniques 3rd ed*. Massachusetts (US): Morgan Kaufmann Publisher.
- [23] *Peraturan Menteri Kehutanan Nomor: P. 12/Menhut-I/2009 tentang Pengendalian Kebakaran Hutan*. (2009). <https://doi.org/10.18860/ling.v5i1.609>
- [24] Pramesti, D. F., Furqon, M. T., & Dewi, C. (2017). *Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan / Lahan Berdasarkan Persebaran Titik Panas (Hotspot)*. 1(9), 723–732.
- [25] Putrada, A. M., Abdurohman, M., & Putrada, A. G. (2018). Increasing Smoke Classifier Accuracy using Naïve Bayes Method on Internet of Things. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 4(1), 19–26. <https://doi.org/10.22219/kinetik.v4i1.704>
- [26] Reza Noviansyah, M., Rismawan, T., Marisa Midyanti, D., Sistem Komputer, J., & MIPA Universitas Tanjungpura Jl Hadari Nawawi, F. H. (2018). Penerapan Data Mining Menggunakan Metode K-Nearest Neighbor Untuk Klasifikasi Indeks Cuaca Kebakaran Berdasarkan Data Aws (Automatic Weather Station) (Studi Kasus: Kabupaten Kubu Raya). *Jurnal Coding, Sistem Komputer Untan*, 6(2), 48–56. Retrieved from <http://jurnal.untan.ac.id/index.php/jcskommipa/article/view/26672>
- [27] Rizka, Y. A. (2016). *Klasifikasi kemunculan titik panas pada lahan gambut di Sumatera dan Kalimantan menggunakan algoritme naïve bayes*.
- [28] Roswintarti, D. O. (2016). *Informasi Titik Panas (Hotspot) Kebakaran Hutan/Lahan. 1*.
- [29] Saleh, A. (2015). *Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga*. 2(3), 207–217.
- [30] Setiawan, A. (2018). Kebijakan Penanganan Kebakaran Hutan dan Lahan di Indonesia. *Political Ecology*, VI(May), 1–13.
- [31] Suntoro, J. (2019). *Data Mining: Algoritma dan Implementasi dengan Pemrograman PHP*. Jakarta: PT ELEX MEDIA KOMPUTINDO.
- [32] Widiastuti, D., Informasi, J. S., & Gunadarma, U. (2012). *ANALISA PERBANDINGAN ALGORITMA SVM , NAIVE BAYES , DAN DECISION TREE DALAM MENGLASIFIKASIKAN SERANGAN (ATTACKS)*. 1–8.
- [33] Witten, Ian H. Frank, E. H. (2011). Data Mining Practical Machine Learning Tools and Techniques. In *Encyclopedia of Ecology, Five-Volume Set* (third). <https://doi.org/10.1016/B978-008045405-4.00153-1>
- [34] Wu, X., Kumar, V., ... Dan, J. H. (2008). *Top 10 algorithms in data mining*. <https://doi.org/10.1007/s10115-007-0114-2>