

Klasifikasi Kebakaran Hutan Menggunakan Metode *K-Nearest Neighbor* : Studi Kasus Hutan Provinsi Kalimantan Barat

*(Forest Fires Classification by Using The K-Nearest Neighbor
Method: Case Studi of West Kalimantan Province Forest)*

Ari Rudiyan^{[1]*}, Akhmad Erik Dzulkipli^[2], Khabib Munazar^[3]

^{[1],[2],[3]} Teknik Informatika, Universitas Amikom Yogyakarta

E-mail: ari@students.amikom.ac.id, erikdzulkifli@students.amikom.ac.id,
khabibmunazar@students.amikom.ac.id

KEYWORDS:

*Forest Fire, Classification,
K-Nearest Neighbor, West Kalimantan*

ABSTRACT

A very wide forest could cause natural disaster such as forest fire that resulting losses to inhabitant, one of which effecting health and safety. West Kalimantan province is one of the province in Indonesia that has wide area of forest with 8,200,000 ha of forest and 1,600,000 ha of peatland all over the Kalimantan Island. Therefor this study is focusing on the data of west Kalimantan province forest. The aim of the study is to classify forest fire in West Kalimantan Province and followed by designing a REST API application of forest fire detector. In hope that in future, the application will be useful to prevent forest fire in the area of west Kalimantan Province. K-Nearest Neighbor method and balltree algorithm are used in this study to collect and process the data. The sample that are collected about 30% of 14,201 data with accuracy up to 92% with K = 18.

KATA KUNCI:

**Kebakaran Hutan, Klasifikasi,
K-Nearest Neighbor, Kalimantan
Barat**

ABSTRAK

Hutan yang sangat luas memungkinkan timbulnya bencana alam seperti kebakaran hutan yang dapat menimbulkan kerugian-kerugian bagi masyarakat, salah satunya berdampak pada kesehatan dan keselamatan jiwa. Provinsi Kalimantan Barat merupakan salah satu daerah di Indonesia yang memiliki kawasan hutan yang sangat luas, tercatat 8.200.000 ha hutan dan 1.600.000 ha lahan gambut yang terbentang di bagian wilayah pulau Kalimantan, oleh itu peneliti menggunakan data hutan provinsi Kalimantan Barat sebagai objek penelitian. Penelitian ini bertujuan untuk melakukan klasifikasi kebakaran hutan di provinsi Kalimantan Barat yang dilanjutkan dengan perancangan REST API aplikasi deteksi kebakaran hutan. Peneliti menggunakan metode k-Nearest Neighbor dan menggunakan algoritma balltree sebagai teknik pencarian tetangga terdekat. Data testing yang digunakan sejumlah 30% dari 14.201 data, dan data training yang digunakan sejumlah 70% dari 14.201 data. Hasil yang didapatkan berupa tingkat akurasi sebesar 92% dengan nilai K = 18. Diharapkan penelitian ini dapat bermanfaat dalam pencegahan kebakaran hutan di Provinsi Kalimantan Barat.

I. PENDAHULUAN

Provinsi Kalimantan Barat merupakan salah satu daerah di Indonesia yang memiliki kawasan hutan yang sangat luas tercatat 8.200.000 ha hutan dan 1.600.000 ha lahan gambut yang terbentang di bagian wilayah pulau Kalimantan [1]. Hutan yang sangat luas ini tentunya banyak menimbulkan bencana alam seperti kebakaran hutan yang dapat menimbulkan kerugian-kerugian bagi masyarakat, salah satunya berdampak pada kesehatan dan

keselamatan jiwa. Menurut kementerian Kesehatan (2015) terjadi kebakaran hutan dan lahan yang sangat parah di beberapa provinsi di Indonesia yang menyebabkan beberapa negara di Asia Tenggara terkena polusi udara [2]. Dampak yang ditimbulkan tersebut sangat berbahaya oleh karena itu pemerintah terus berusaha melakukan pencegahan untuk mengurangi tingkat kebakaran. Intensitas kebakaran tertinggi pada umumnya terjadi pada musim

kemarau yang mempengaruhi hutan mudah terbakar karena curah hujan sangat rendah [3].

Kebakaran hutan dapat dipengaruhi oleh keadaan cuaca yang meliputi keadaan suhu, kecepatan angin, dan curah hujan [4]. Label prediksi ini dapat dihasilkan melalui proses klasifikasi sehingga dapat digunakan variabel untuk melakukan prediksi terjadinya kebakaran hutan atau tidak. Terdapat beragam metode klasifikasi yang dapat digunakan diantaranya adalah metode *K-Nearest Neighbor* (KNN). Kelebihan menggunakan metode KNN yaitu sangat cepat melakukan pelatihan dengan data besar walaupun data tersebut memiliki derau [5]. Sehingga pada penelitian ini peneliti memutuskan untuk menggunakan KNN sebagai metode klasifikasi.

Penelitian ini bertujuan untuk melakukan klasifikasi kebakaran hutan di provinsi Kalimantan Barat. Penelitian serupa dilakukan oleh [4] di kabupaten Kubu Raya dengan memanfaatkan *Automatic Weather Station* (AWS) yang menghasilkan indeks cuaca kebakaran tingkat rendah, sedang, tinggi, dan ekstrim. Penelitian tersebut diperuntukkan petugas Manggala Agni agar mempermudah mengetahui data kebakaran pada minggu selanjutnya. Berbeda dengan penelitian ini, dimana hasil akhirnya berupa *REST API* prediksi kebakaran hutan yang dapat digunakan oleh masyarakat luas khususnya di provinsi Kalimantan Barat. Hasil prediksi berupa akan terjadi kebakaran atau tidak dengan menampilkan nilai *confidence* rendah (*low*), sedang (*medium*), dan tinggi (*high*). Penelitian ini diharapkan dapat dijadikan sebagai *tools* untuk meningkatkan kewaspadaan dini terhadap kebakaran hutan sehingga dampak-dampak yang diakibatkan oleh kebakaran tersebut dapat dikurangi.

II. METODOLOGI

A. Kajian Pustaka

1) Kebakaran Hutan

Kebakaran hutan merupakan bencana alam yang dapat terjadi kapan saja terutama pada musim kemarau, kebakaran ini dapat disebabkan karena ulah manusia ataupun faktor bawaan alam yaitu cuaca [3]. Dampak kebakaran hutan sangat beragam, mulai dari masalah ekonomi, sosial, dan kesehatan. Penelitian sebelumnya telah dilakukan Maharani (2020) mengenai dampak kabut asap yang ditimbulkan oleh kebakaran hutan di kota

Jambi. Hasil penelitian tersebut menggunakan 100 orang terdampak sebagai sample dan menghasilkan 56% mengalami gejala batuk, 52% iritasi mata, 47% sesak nafas, 33% iritasi hidung, 16% nyeri tenggorokan dan 2% mengalami diare [6].

Penelitian dibidang Informatika telah dilakukan oleh Pratiwi, dkk (2021) tentang Klasifikasi Kebakaran Hutan menggunakan *Algoritma Naive Bayes* di Kabupaten Pelalawan. Klasifikasi dilakukan menggunakan atribut suhu, kelembaban, curah hujan dan kecepatann angin. Tingkat akurasi tertinggi yang dihasilkan adalah 81,03% dengan menggunakan dataset pada tahun 2017 [3].

Selanjutnya penelitian juga telah dilakukan Tri Basuki Kurniawan (2021) dengan judul Prediksi Lokasi Titik Panas Kebakaran Hutan Menggunakan Model Regresion SVM (*Support Vector Machine*) pada Data Kebakaran Hutan Daops Manggala Agni Oki Provinsi Sumatera Selatan Tahun 2019. Tujuan penelitian ini yaitu memprediksi lokasi titik hotspot kebakaran hutan yang dapat digunakan untuk melakukan prediksi kebakaran hutan di masa yang akan datang. Penelitian ini menggunakan data hotspot pada tahun 2019 yang digunakan untuk melakukan prediksi di tahun 2020, dan untuk prediksi tahun 2021 digunakan data pada tahun 2019 dan 2020, begitupun seterusnya. Hasil penelitian ini menunjukkan pada tahun 2021 terdapat titik hotspot terbanyak di kecamatan Pematang Panggal dengan jumlah 448 lokasi hotspot. Disusul pada tahun 2022 diprediksi 571 titik lokasi hotspot terbanyak pada kecamatan Cengal [7].

Selanjutnya penelitian sebelumnya dengan menggunakan metode yang sama telah dilakukan oleh Reza Noviansyah (2018), judul yang diangkat yaitu Penerapan Data Mining Menggunakan Metode *K-Nearest Neighbor* untuk Klasifikasi Indeks Cuaca Kebakaran berdasarkan Data AWS (*Automatic Weather Station*) (Studi Kasus: Kabupaten Kubu Raya). Penelitian ini bertujuan untuk membuat system guna mempermudah petugas Manggala Agni untuk mengetahui grafik data kebakaran pada minggu yang akan datang. Hasil penelitian ini mendapatkan persentase akurasi keberhasilan sebesar 80,16% dengan nilai $K = 5$, pada percobaan ini menggunakan 252 data uji [4].

Penelitian selanjutnya pernah dilakukan oleh Eka Bagus Susanto (2019) mengenai Analisis

Perbandingan *Algoritma Naïve Bayes* dan *K-Nearest Neighbor* untuk Klasifikasi Multi Dataset. Penelitian tersebut memperoleh hasil rata-rata keakurasian sejumlah 78,38% menggunakan metode *Naïve Bayes* dan 93,17% menggunakan KNN [8].

2) Estimasi

Estimasi adalah suatu proses yang dilakukan guna memprediksi atau memperkirakan sesuatu yang belum diketahui nilainya secara pasti.

3) Prediksi

Prediksi adalah suatu proses peramalan dimasa yang akan datang, prediksi dapat dilakukan berdasarkan data yang terdapat pada masa lampau.

4) Klasifikasi

Klasifikasi adalah suatu kegiatan yang dilakukan untuk mendapatkan suatu model yang dapat membedakan kelas data yang selanjutnya digunakan untuk kepentingan tertentu. Aktifitas klasifikasi ini adalah cara untuk menemukan label yang tidak diketahui sebelumnya pada suatu kelas. Label yang diperoleh tersebut dapat digunakan untuk melakukan prediksi [3]. Ada banyak metode klasifikasi yang dapat digunakan salah satunya adalah *K-Nearest Neighbor (KNN)*.

5) K-Nearest Neighbor

Penelitian ini menggunakan metode klasifikasi *K-Nearest Neighbor (KNN)*. KNN merupakan metode klasifikasi tertua dan populer, metode ini bekerja dengan mempertimbangkan kelas terdekat dari suatu object berdasarkan nilai K [9]. Jarak tersebut ditentukan dengan cara melakukan training dan testing terlebih dahulu. Selanjutnya dilakukan perhitungan jarak dari setiap data testing terhadap data training (Jarak euclidean) [10]. Berikut untuk menghitung jarak tersebut dapat digunakan rumus sebagai berikut.

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad [1]$$

Keterangan :

$D(x_i, x_j)$: Jarak Eueclidience
x_i, x_j	: record ke-m, record ke-n
(a_r)	: data ke-r
m, n	: 1, 2, 3, ... n
n	: dimensi objek

Terdapat beberapa jenis algoritma pencarian tetangga terdedat seperti pemindaian linier, pohon KD, pohon balltree, pohon metrik, dan *hash sensitif lokal (LSH)*.

6) Titik Panas (Hotspot)

Titik panas (*hotspot*) dapat dijadikan sebagai indikator kebakaran hutan yang mendeteksi suatu lokasi yang memiliki suhu relatif lebih tinggi dibandingkan dengan suhu disekitarnya. Titik panas dapat dideteksi menggunakan sensor, sensor tersebut merepresentasikan suhu titik panas berdasarkan *brighness temperature*. *Brighness temperature* adalah salah satu fitur yang dimiliki pada citra penginderaan jarak jauh yang dapat dideteksi dari lokasi tertentu.

B. Pengumpulan Data

Pengumpulan data dilakukan dengan cara mencari data publik diberbagai sumber terpercaya yang pada nantinya dapat diolah menjadi dataset siap pakai. Data yang digunakan dalam penelitian ini yaitu data publik yang didapatkan melalui situs Kementerian Lingkungan Hidup dan Kehutanan Republik Indonesia (<https://sipongi.menlhk.go.id/>). Data yang digunakan adalah data yang terekam sejak Bulan Agustus 2021 sampai dengan Bulan September 2021 data tersebut berupa data Stasiun Meteorologi yang tersebar diberbagai kabupaten di Provinsi Kalimantan Barat, data titik hotspot, dan keadaan cuaca. Ketiga data tersebut didapatkan secara terpisah dan butuh dilakukan *pre-processing data* sehingga dapat digunakan untuk melakukan prediksi.

C. Pre-Processing Data

Tahap *pre-processing data* pada penelitian ini dilakukan menggunakan *Google colab* dengan bantuan *library pandas*. Ada 3 (tiga) data yang diolah untuk dijadikan sebagai dataset, yaitu data kebakaran yang telah diperoleh selama proses pengumpulan data seperti stasiun Meteorologi, keadaan cuaca, dan titik hotspot kebakaran di provinsi Kalimantan Barat.

D. Training dan Testing

Tahap pembelajaran (*training*) dan pengujian (*testing*) dilakukan setelah selesai melakukan *pre-processing data*. Data tersebut dibagi menjadi dua bagian yaitu 70% digunakan untuk melakukan

training dan 30% digunakan untuk melakukan testing menggunakan metode KNN. Pada proses ini peneliti sudah dapat memperoleh hasil berupa tingkat akurasi yang dihasilkan.

Latitude_rad	Longitude_rad	SUHU	CURAHHUJUAN	LAMAPENYINARAN	KECEPATANANGIN
732	0.017899	1.914303	2	0	2
8246	0.017057	1.971066	2	0	1
1950	-0.004967	1.939707	2	0	1
148	0.034113	1.910977	2	0	2
4342	0.005807	1.933463	2	0	1
...
11262	0.013566	1.909904	2	1	2
1010	-0.008028	1.931645	2	0	1
10338	-0.009328	1.922661	2	0	2
11822	0.015289	1.926142	1	3	1
3644	0.012546	1.935772	2	0	1

Gbr. 1 Gambar Tabel Data Testing

Latitude_rad	Longitude_rad	SUHU	CURAHHUJUAN	LAMAPENYINARAN	KECEPATANANGIN
1117	-0.009922	1.927269	2	0	2
9034	0.002188	1.924905	2	1	2
5397	0.002430	1.935030	2	0	2
3579	0.014789	1.941323	2	0	1
12457	0.007561	1.915804	2	0	2
...
12786	-0.014869	1.925951	2	0	2
13634	-0.040735	1.924415	2	0	1
5590	0.003487	1.927918	2	0	2
8934	0.010617	1.926895	2	1	2
711	0.016739	1.912926	2	0	2

Gbr. 2 Gambar Tabel Data Training

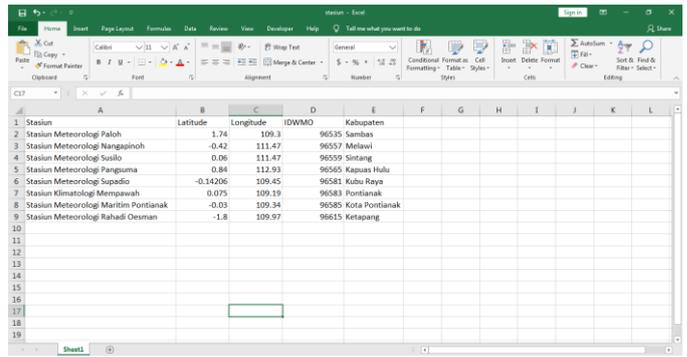
E. REST API

Setelah berhasil melakukan testing selanjutnya dilakukan proses perancangan REST API dalam bentuk website. REST API ini diharapkan dapat digunakan dan dikembangkan oleh pihak lain untuk melakukan prediksi kebakaran hutan di provinsi Kalimantan barat

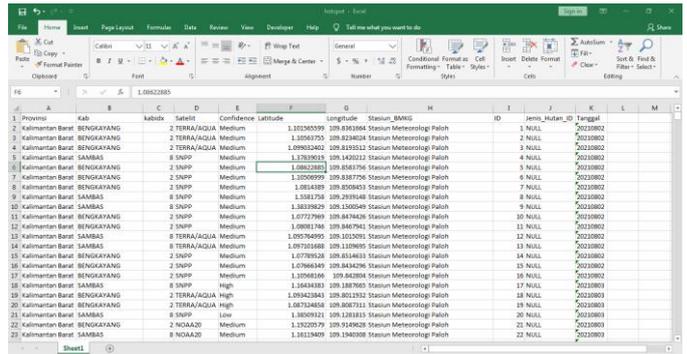
III. HASIL DAN PEMBAHASAN

A. Data Kebakaran Hutan

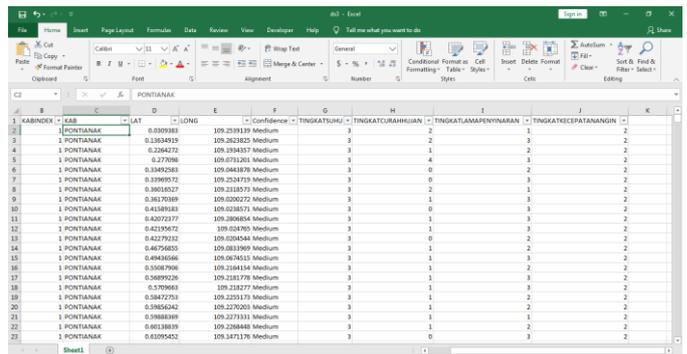
Data kebakaran hutan diperoleh dari situs Kementerian Kehutanan dan Lingkungan hidup Indonesia <http://sipongi.menlhk.go.id/>. Data-data yang diperoleh merupakan data mentah yang selanjutnya diperlukan proses pengolahan data terlebih dahulu agar dapat dipergunakan untuk prediksi. Data tersebut dapat dilihat pada gambar berikut.



Gbr. 3 Stasiun Meteorologi Provinsi Kalimantan Barat Dalam Format .xlsx



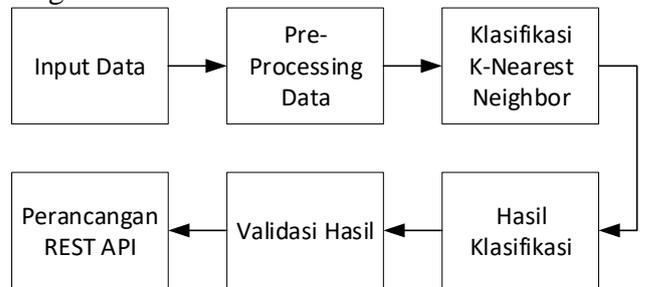
Gbr. 4 Sebaran Titik Hotspot Kebakaran Hutan Provinsi Kalimantan Barat Dalam Format .xlsx



Gbr. 5 Keadaan Cuaca Dalam Format .xlsx

B. Proses Perancangan

Proses Perancangan pada penelitian ini diawali dengan membuat alur kerja sistem yang dapat dilihat pada gambar 6 berikut.



Gbr. 6 Diagram Blok Proses

Seluruh proses ini dilakukan menggunakan pemrograman python 3 dengan bantuan *google colab* dan *library pandas*.

1) Input Data

Proses pertama yang dilakukan adalah input data kebakaran hutan yang diperoleh sesuai dengan poin 1 diatas. File yang diinput berupa file excel diantaranya adalah *hotspot.xlsx*, *stasiun.xlsx*, dan *cuaca.xlsx*. *Source code* yang digunakan pada proses ini dapat dilihat pada gambar 5 berikut.

```
[ ] from google.colab import drive
import matplotlib.pyplot as plt
import pandas as pd
from pandas.io.json import json_normalize
drive.mount('/content/drive')

Mounted at /content/drive

[ ] import numpy as np
from sklearn.neighbors import BallTree, KDTree

Klik dua kali (atau tekan Enter) untuk mengedit

datahotspot = pd.read_excel ('/content/drive/MyDrive/Colab Notebooks/Hutan/hotspot.xlsx')
datahotspot.head()
datastasiun = pd.read_excel ('/content/drive/MyDrive/Colab Notebooks/Hutan/stasiun.xlsx')
datastasiun.head()
datacuaca = pd.read_excel ('/content/drive/MyDrive/Colab Notebooks/Hutan/cuaca.xlsx')
datacuaca.head()
len(datahotspot)
```

	Stasiun	Latitude	Longitude	ID#NO	Kabupaten
0	Stasiun Meteorologi Paloh	1.74000	109.30	96535	Sambas
1	Stasiun Meteorologi Nangapinoh	-0.42000	111.47	96557	Melawi
2	Stasiun Meteorologi Susilo	0.06000	111.47	96559	Sintang
3	Stasiun Meteorologi Pangsuma	0.84000	112.93	96565	Kapuas Hulu
4	Stasiun Meteorologi Supadio	-0.14206	109.45	96581	Kubu Raya

Gbr. 7 Source Code Input Data

2) Pre-Processing Data

Selanjutnya dilakukan *pre-processing data* yang telah diinput agar dapat digunakan sebagai data training dan data testing. Tahap ini dilakukan pengubahan data pada atribut koordinat *Latitude* dan *Longitude* menjadi radian. Proses ini bertujuan untuk :

- Untuk pengambilan data cuaca, hotspot kebakaran hutan , dan stasiun
- Mengubah nilai *latitude* dan *longitude* menjadi ke radian.
- Menambahkan stasiun ID pada hotspot sesuai jarak *latitude* dan *longitude*
- Menambahkan atribut cuaca pada stasiun ID
- Mengelompokkan dan melakukan tranformasi data dari tipe data numerik ke data nominal, pengelompokan tersebut dapat dilihat pada table 1 berikut:

TABEL I
TRANSFORMASI DATA NUMERIK

Atribut	Data Numerik	Data Nominal
Suhu	0 °C – 10 °C	Rendah
	11 °C – 26 °C	Sedang
	27 °C – 60 °C	Tinggi
	>60 °C	Sangat Tinggi
Curah Hujan	0 – 0.5 mm/hari	Tidak Hujan
	0.6 – 20 mm/hari	Rendah
	21 – 50 mm/hari	Sedang
	>50 mm/hari	Tinggi
Lama Penyinaran	0 – 2 jam	Rendah
	3 – 5 jam	Sedang
	6 – 12 jam	Tinggi
Kecepatan Angin	0 – 1	Rendah
	1 – 5	Sedang
	6 – 11	Tinggi
	12 – 19	Sangat Tinggi

Hasil *pre-processing* yang telah dilakukan diatas dapat dilihat pada gambar berikut.

Provinsi	Kab	cabids	Satelit	Confidence	Latitude	Longitude	Stasiun_BMKG	ID	Tanggal
Kalimantan Barat	BENGKAYANG	2	TERRA/AQUA	Medium	1.101565599	109.8361664	Stasiun Meteorologi Paloh	1	20210802
Kalimantan Barat	BENGKAYANG	2	TERRA/AQUA	Medium	1.10563755	109.8234024	Stasiun Meteorologi Paloh	2	20210802
Kalimantan Barat	BENGKAYANG	2	TERRA/AQUA	Medium	1.099032402	109.8193512	Stasiun Meteorologi Paloh	3	20210802
Kalimantan Barat	SAMBAS	8	SNPP	Medium	1.37893019	109.1420212	Stasiun Meteorologi Paloh	4	20210802
Kalimantan Barat	BENGKAYANG	2	SNPP	Medium	1.08622885	109.8537794	Stasiun Meteorologi Paloh	5	20210802
Kalimantan Barat	BENGKAYANG	2	SNPP	Medium	1.10506999	109.8387756	Stasiun Meteorologi Paloh	6	20210802
Kalimantan Barat	BENGKAYANG	2	SNPP	Medium	1.0814389	109.8508453	Stasiun Meteorologi Paloh	7	20210802
Kalimantan Barat	SAMBAS	8	SNPP	Medium	1.5581758	109.2939148	Stasiun Meteorologi Paloh	8	20210802
Kalimantan Barat	SAMBAS	8	SNPP	Medium	1.3839829	109.1500549	Stasiun Meteorologi Paloh	9	20210802
Kalimantan Barat	BENGKAYANG	2	SNPP	Medium	1.07727969	109.8474426	Stasiun Meteorologi Paloh	10	20210802
Kalimantan Barat	BENGKAYANG	2	SNPP	Medium	1.08081746	109.8647941	Stasiun Meteorologi Paloh	11	20210802
Kalimantan Barat	SAMBAS	8	TERRA/AQUA	Medium	1.095764995	109.1015091	Stasiun Meteorologi Paloh	12	20210802
Kalimantan Barat	SAMBAS	8	TERRA/AQUA	Medium	1.097101688	109.1109695	Stasiun Meteorologi Paloh	13	20210802
Kalimantan Barat	BENGKAYANG	2	SNPP	Medium	1.07789528	109.8514633	Stasiun Meteorologi Paloh	14	20210802
Kalimantan Barat	BENGKAYANG	2	SNPP	Medium	1.07666349	109.8434296	Stasiun Meteorologi Paloh	15	20210802
Kalimantan Barat	BENGKAYANG	2	SNPP	Medium	1.10568166	109.842804	Stasiun Meteorologi Paloh	16	20210802
Kalimantan Barat	SAMBAS	8	SNPP	High	1.16543883	109.1867655	Stasiun Meteorologi Paloh	17	20210803
Kalimantan Barat	BENGKAYANG	2	TERRA/AQUA	High	1.093423843	109.8011932	Stasiun Meteorologi Paloh	18	20210803
Kalimantan Barat	BENGKAYANG	2	TERRA/AQUA	High	1.087324858	109.8087311	Stasiun Meteorologi Paloh	19	20210803
Kalimantan Barat	SAMBAS	8	SNPP	Low	1.38509321	109.1281815	Stasiun Meteorologi Paloh	20	20210803
Kalimantan Barat	BENGKAYANG	2	NOAA20	Medium	1.19220579	109.9149678	Stasiun Meteorologi Paloh	21	20210803

Gbr. 8 Data Hotspot Sebelum Pre-Processing

	cabids	Confidence	Latitude_rad	Longitude_rad	SUHU	CURAHHUJUAN	LAMAPENYINARAN	KECEPATANANGIN
0	2	Medium	0.019225947	1.917002741	2	0	1	
1	2	Medium	0.019297016	1.916779968	2	0	1	
2	2	Medium	0.019181734	1.916709261	2	0	1	
3	8	Medium	0.024057447	1.904887622	2	0	1	
4	2	Medium	0.01895827	1.917390364	2	0	1	
5	2	Medium	0.01928711	1.917048281	2	0	1	
6	2	Medium	0.018874669	1.917258937	2	0	1	
7	8	Medium	0.027195298	1.907538665	2	0	1	
8	8	Medium	0.024144855	1.905027837	2	0	1	
9	2	Medium	0.018802078	1.917199549	2	0	1	
10	2	Medium	0.018863823	1.91718823	2	0	1	
11	8	Medium	0.019124707	1.904180553	2	0	1	
12	8	Medium	0.019148037	1.904345669	2	0	1	
13	2	Medium	0.018812822	1.917269723	2	0	1	
14	2	Medium	0.018791323	1.917129508	2	0	1	
15	2	Medium	0.019297785	1.917118589	2	0	1	
16	8	High	0.020321633	1.905703481	2	0	1	
17	2	High	0.019083846	1.916392345	2	0	1	
18	2	High	0.018977399	1.916523905	2	0	1	
19	8	Low	0.024174437	1.904646073	2	0	1	
20	2	Medium	0.020807916	1.918377998	2	0	1	
21	8	Medium	0.02026666	1.90579536	2	0	1	
22	2	Medium	0.018754725	1.917318459	2	0	1	
23	2	Medium	0.015522283	1.917121385	2	0	1	
24	13	Medium	0.016594649	1.91937402	2	0	1	

Gbr. 9 Data Hotspot Setelah Pre-processing

Pada proses ini dilakukan data *cleansing* seperti atribut tanggal dan ID dihilangkan. Selanjutnya didapatkan hasil berupa atribut baru seperti tampak

pada gambar 10. *Source code* yang digunakan pada proses ini dapat dilihat pada gambar berikut.

```
[ ] # Creates new columns converting coordinate degrees to radians.
for column in datahotspot[["Latitude", "Longitude"]]:
    rad = np.deg2rad(datahotspot[column].values)
    datahotspot[f'{column}_rad'] = rad
for column in datastasiun[["Latitude", "Longitude"]]:
    rad = np.deg2rad(datastasiun[column].values)
    datastasiun[f'{column}_rad'] = rad
```

Gbr. 10 *Source Code* Koordinat ke Radian

3) *Klasifikasi K-Nearest Neighbor*

Selanjutnya pada tahap ini dilakukan pengklasifikasian data menggunakan algoritma *K-Nearest Neighbor* (KNN). Proses yang dilakukan menggunakan algoritma ini dapat dilihat pada gambar 9 berikut.



Gbr. 11 *Flowchart K-Nearest Neighbor*

Penjelasan untuk *flowchart* diatas dapat dijabarkan sebagai berikut.

- Memulai proses
- Masukkan Data Uji
Data yang dimasukkan pada proses ini berupa suhu, kecepatan angin, lama penyinaran, dan curah hujan.
- Tentukan Nilai K
Menentukan nilai K terdekat untuk mendapatkan hasil berupa nilai akurasi. Perolehan nilai akurasi akan berpengaruh sesuai jumlah K yang dimasukkan.

- Hitung jarak data uji ke data latih
Proses penghitungan jarak ini ditentukan menggunakan rumus *euclidean distance*.
- Urutkan Hasil Perhitungan Jarak
Nilai jarak yang telah didapatkan menggunakan rumus *euclidean distance* kemudian dilakukan proses pengurutan.
- Hasil Klasifikasi
Hasil klasifikasi selanjutnya dapat disimpan untuk dipergunakan sebagai *database*.
- End (Selesai)

Source code untuk mengeksekusi algoritma ini dapat dilihat pada Gbr. 10 berikut.

```
[ ] from sklearn.neighbors import KNeighborsClassifier
from sklearn import metrics

knn = KNeighborsClassifier(n_neighbors=2, weights='distance', algorithm='ball_tree')

knn.fit(DataX, DataY)
y_pred = knn.predict(X_test)
score = metrics.accuracy_score(y_test, y_pred)
print("Akurasi dengan menggunakan Nearest Neighbor: ", score)
hasil = X_test.copy()
h=[]
for i in range(len(y_pred)):
    if y_pred[i]==y_test.iloc[i].Confidence :
        h.append('Benar')
    else:
        h.append('Salah')

for i in range(len(X_test)):
    hasil[f'prediksi'] = y_pred
    hasil[f'label'] = y_test
    hasil[f'hasil'] = h
    hasil.to_csv('sample_data/hasiltest.csv', index = True)

/usr/local/lib/python3.7/dist-packages/sklearn/neighbors/_classification.py:198: DataConversionWarning:
return self._fit(X, y)
Akurasi dengan menggunakan Nearest Neighbor: 1.0
```

Gbr. 12 *Source Code* Klasifikasi KNN

4) *Hasil Klasifikasi*

Hasil klasifikasi dengan data training sejumlah 14201 *record data* dengan 3 kelas target (*low, medium dan high*). Selanjutnya diperoleh hasil testing dengan pembagian data sejumlah 30% didapatkan 92% tingkat akurasi dengan menggunakan nilai K = 18.

5) *Validasi Hasil*

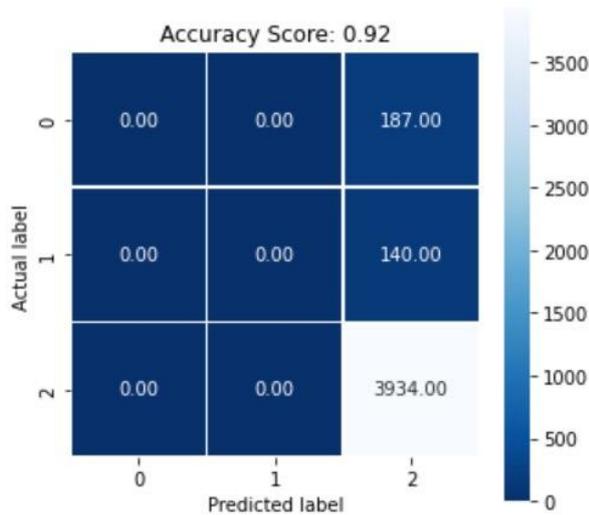
Proses validasi ini dilakukan dengan mencocokkan hasil prediksi dengan nilai sebenarnya. Validasi dilakukan menggunakan teknik *grid search cross validation* (*grindseachCV*) dan mendapatkan hasil sesuai pada table II berikut.

TABEL II
HASIL VALIDASI

Jumlah K	Akurasi
1	0.814286
2	0.903119
3	0.897686
4	0.913280
5	0.913883
6	0.918410
7	0.918813

Jumlah K	Akurasi
8	0.920121
9	0.920121
10	0.921429
11	0.921227
12	0.921831
13	0.922334
14	0.922535
15	0.922535
16	0.922636
17	0.922535
18	0.922736
...	...
28	0.922736

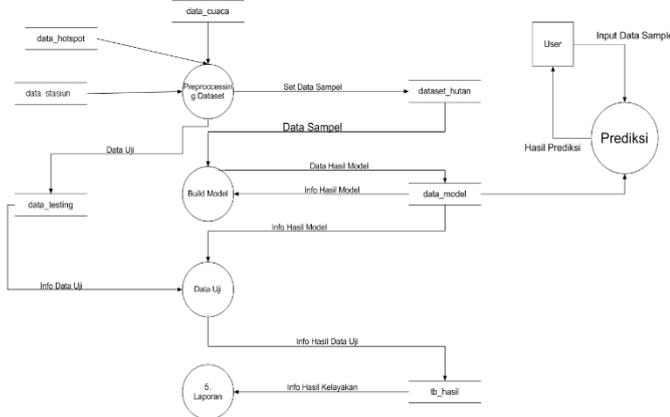
Visualisasi menggunakan *confusion matrix* untuk mengetahui tingkat akurasi nilai K = 18 dapat dilihat pada gambar berikut.



Gbr. 13 Validasi Hasil

C. Data Flow Diagram

Data Flow Diagram dapat digambarkan pada gambar berikut :



Gbr. 14 Data Flow Diagram Level 1

Berdasarkan serangkaian proses yang digambarkan oleh Data Flow Diagram (DFD) diatas diketahui bahwa user hanya perlu memasukkan data sample untuk melakukan prediksi. Data sample tersebut berupa suhu, curah hujra, kecepatan angin, dan lama penyinaran. Setelah itu *user* akan mendapatkan umpan balik berupa hasil prediksi.

D. REST API Forest Fire Detection System

Setelah berhasil melakukan training dan testing langkah selanjutnya pada penelitian ini yaitu tahap pembuatan REST API aplikasi Deteksi Kebakaran Hutan berbasis *website*. REST API ini nantinya dapat digunakan sebagai *tools* untuk melakukan prediksi kebakaran hutan di provinsi Kalimantan Barat dengan hanya memasukkan atribut sesuai yang dibutuhkan. Data tersebut tinggal dimasukkan ke dalam *field* tersedia, seperti *Latitude*, *Longitude*, suhu udara dalam derajat *celcius*, curah hujan, lama penyinaran, dan kecepatan angin. Selain itu REST API ini dapat digunakan sebagai bahan pembelajaran rumpun mata kuliah terkait dan juga dapat dikembangkan menjadi aplikasi yang lebih baik.. REST API aplikasi tersebut dapat dilihat pada Gbr. 13 berikut.

Gbr. 15 Halaman REST API Simulator

REST API diatas dilakukan pengujian dengan memasukkan *latitude* 1.16434383, *longitude* 109.18876648, suhu udara 28 °C, curah hujan 0.5 mm/hari, lama penyinaran 5 jam, kecepatan angin rata-rata 2 didapatkan hasil prediksi dengan nilai *confidence High* yang artinya potensi terjadinya kebakaran tinggi.

IV. KESIMPULAN

Serangkaian proses klasifikasi dengan metode yang telah ditentukan yaitu menggunakan *K-Nearest Neighbor* dan menggunakan *balltree*

sebagai teknik pencarian terdekat. Berdasarkan proses tersebut didapatkan hasil pengujian menggunakan data testing sejumlah 30% dari 14.201 data, dan data *training* yang digunakan sejumlah 70% dari 14.201 data, didapatkan tingkat akurasi sebesar **92%** dengan nilai $K = 18$. Nilai yang sama diperoleh dengan jumlah K sampai dengan 28.

Penelitian ini dilakukan sampai dengan pembuatan *REST API* aplikasi prediksi kebakaran hutan, oleh karena itu disarankan kepada peneliti selanjutnya untuk menjadikan hasil akhir yang telah dibuat ini terintegrasi dengan berbagai API website penyedia data sample yang dibutuhkan.

UCAPAN TERIMAKASIH

Puji syukur kami panjatkan kepada Allah SWT. yang telah memberikan peneliti kesehatan dan kesempatan untuk menyelesaikan penelitian ini. Terimakasih yang setinggi-tingginya kami sampaikan kepada dosen dan tutor pengampu mata kuliah *Data Science* Universitas Amikom Yogyakarta atas ilmu dan bimbingannya selama proses pembelajaran dan penelitian ini dilakukan.

REFERENSI

- [1] “Kalimantan Barat,” *INCAS*. <http://incas.menlhk.go.id/data/west-kalimantan/> (diakses Des 20, 2021).
- [2] A. Yusuf, Hapsoh, S. H. Siregar, dan D. R. Nurrochmat, “Analisis Kebakaran Hutan Dan Lahan Di Provinsi Riau | Yusuf | *Dinamika Lingkungan Indonesia*.” <https://dli.ejournal.unri.ac.id/index.php/DL/article/view/7457/0> (diakses Des 20, 2021).
- [3] T. A. Pratiwi, M. Irsyad, R. Kurniawan, S. Agustian, dan B. S. Negara, “Klasifikasi Kebakaran Hutan Dan Lahan Menggunakan Algoritma Naïve Bayes Di Kabupaten Pelalawan,” *CESS (Journal of Computer Engineering, System and Science)*, vol. 6, no. 1, Art. no. 1, Feb 2021, doi: 10.24114/cess.v6i1.22555.
- [4] M. R. Noviansyah, T. Rismawan, dan D. M. Midyanti, “PENERAPAN DATA MINING MENGGUNAKAN METODE K-NEAREST NEIGHBOR UNTUK KLASIFIKASI INDEKS CUACA KEBAKARAN BERDASARKAN DATA AWS (AUTOMATIC WEATHER STATION) (STUDI KASUS: KABUPATEN KUBU RAYA),” *Coding Jurnal Komputer dan Aplikasi*, vol. 6, no. 2, Art. no. 2, Jul 2018, doi: 10.26418/coding.v6i2.26672.
- [5] S. Mutfrofin, A. Izzah, A. Kurniawardhani, dan M. Masrur, “OPTIMASI TEKNIK KLASIFIKASI MODIFIED K NEAREST NEIGHBOR MENGGUNAKAN ALGORITMA GENETIKA,” *Jurnal Gamma*, vol. 10, no. 1, Art. no. 1, 2014, Diakses: Des 20, 2021. [Daring]. Tersedia pada: <https://ejournal.umm.ac.id/index.php/gamma/article/view/2493>
- [6] C. Maharani, A. A. Asrial, B. A. Purba, dan M. Miftahurrahmah, “EDUKASI DAMPAK KESEHATAN DAN UPAYA PERLINDUNGAN DIRI DARI BENCANA KABUT ASAP,” *Medical Dedication (medic) : Jurnal Pengabdian kepada Masyarakat FKIK UNJA*, vol. 3, no. 1, Art. no. 1, Apr 2020, doi: 10.22437/medicaldedication.v3i1.8631.
- [7] J. Yandi, T. B. Kurniawan, E. S. Negara, dan M. Akbar, “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 | Yandi | *InfoTekJar : Jurnal Nasional Informatika dan Teknologi Jaringan*.” <https://jurnal.uisu.ac.id/index.php/infotekjar/article/view/4101> (diakses Des 20, 2021).
- [8] E. BAGUS SUSANTO, “ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES DAN K-NEAREST NEIGHBOR UNTUK KLASIFIKASI MULTI DATASET,” doctoral, Universitas Muhammadiyah Jember, 2019. Diakses: Des 20, 2021. [Daring]. Tersedia pada: <http://repository.unmuhjember.ac.id/7131/>
- [9] T. Y. Prahudaya dan A. Harjoko, “METODE KLASIFIKASI MUTU JAMBU BIJI MENGGUNAKAN KNN BERDASARKAN FITUR WARNA DAN TEKSTUR | Prahudaya | *Jurnal Teknosains*.” <https://jurnal.ugm.ac.id/teknosains/article/view/26972> (diakses Des 20, 2021).
- [10] Y. A. Setianto, K. Kusriani, dan H. Henderi, “Penerapan Algoritma K-Nearest Neighbour Dalam Menentukan Pembinaan Koperasi Kabupaten Kotawaringin Timur,” *Creative Information Technology Journal*, vol. 5, no. 3, Art. no. 3, Sep 2019, doi: 10.24076/citec.2018v5i3.179.