

Penerapan Algoritma K-Means Untuk Pemetaan Biodiversity Kayu Bulat Di Indonesia

Application of the K-Means Algorithm for Mapping Roundwood Biodiversity in Indonesia

Sugi Harsono^{*1}, Tutut Dwi Prihatin², Anwar Sadad³, Kusri⁴, Dina Maulina⁵

^{1,2,3,4,5} Magister Teknik Informatika, Universitas Amikom Yogyakarta

e-mail: ¹sugiharsono85@students.amikom.ac.id, ²wiwi_21552139@students.amikom.ac.id,
³anwar@students.amikom.ac.id, ⁴kusri@amikom.ac.id, ⁵dina.m@amikom.ac.id

Abstrak

Indonesia merupakan salah satu negara di dunia yang memiliki kawasan hutan luas yang tersebar di pulau Sumatera, Jawa, Kalimantan, Sulawesi, Bali, Nusa Tenggara, Maluku dan Papua. Hutan memiliki fungsi sebagai penghasil oksigen dan habitat terbesar keanekaragaman hayati di Dunia. Kebakaran hutan terjadi hampir di seluruh wilayah Indonesia dan merupakan permasalahan yang terus berulang, dan intensitas akan mulai meningkat pada musim kemarau. Kebakaran hutan akan menimbulkan banyak sekali kerugian baik dari segi kesehatan yang akan mengancam keselamatan jiwa maupun material, mempengaruhi kualitas udara, lahan, air, kerusakan fasilitas dan tempat hidup flora dan fauna yang ada, berkurangnya produksi oksigen hingga musnahnya keanekaragaman hayati. Untuk dapat memperkecil kemungkinan terjadinya kebakaran hutan yang terus meningkat diperlukan analisis pemetaan titik penyebaran api, penyebab kebakaran, dan keanekaragaman hayati hutan. Penelitian ini menggunakan metode pengumpulan data primer dataset persebaran titik kebakaran hutan, data statistik kehutanan, kemudian dilanjutkan dengan pengolahan data Algoritma K-Means dan visualisasi pengolahan data menggunakan bahasa pemrograman Python dan mapping menggunakan GeoPandas. Hasil yang didapatkan adalah jenis produksi hutan tertinggi terdapat di Pulau Sumatera dengan jenis Kayu Akasia sebesar 28,573,433.84 m³, kemudian Pulau Jawa jenis Kayu Sengon sebesar 5,431,029.46 m³ dan Pulau Kalimantan jenis Kayu Rimba Campuran. Sedangkan luas area kebakaran terluas pada kawasan hutan terdapat di Pulau Kalimantan sebesar 243,013.5 hektar dari jumlah area hutan Kalimantan 53,054,900 hektar atau sebesar 0.458 %. K-Means 5 klaster antara Pulau Sumatera dan Pulau Jawa adalah Klaster 1 terdapat 82 jenis Kayu, Klaster kedua Kayu Akasia, Klaster ketiga Kayu Ekaliptus, Klaster keempat Kayu Sengon dan Klaster kelima adalah Kayu Rimba Campuran, Mahoni dan Jati.

Kata kunci—kebakaran, k-means, biodiversity, geopandas, hutan

Abstract

Indonesia is one of the countries in the world that has a large forest area spread across the islands of Sumatra, Jawa, Kalimantan, Sulawesi, Bali, Nusa Tenggara, Maluku and Papua. Forests have a function as the largest producer of oxygen and habitat of biodiversity in the World. Forest fires occur almost throughout Indonesia and are a recurring problem, and the intensity will begin to increase in the dry season. Forest fires will cause a lot of losses in terms of health that will threaten life and material safety, affect air quality, land, water, damage facilities and places of life of existing flora and fauna, and reduce oxygen production to destruction of biodiversity. To be able to minimize the possibility of forest fires that continue to increase, it is necessary to analyze the mapping of fire spread points, fire causes, and forest biodiversity. This study used the method of collecting primary data from datasets for the distribution of forest fire points, forestry statistical

data, then continued with data processing of the K-Means Algorithm and visualization of data processing using python programming language and mapping using GeoPandas. The results obtained are the highest type of forest production found on the island of Sumatra with the type of Acacia Wood of 28,573,433.84 m³, then Jawa Island of Sengon Wood type of 5,431,029.46 m³ and Kalimantan Island of Mixed Jungle Wood type. Meanwhile, the largest fire area in the forest area is found on the island of Kalimantan of 243,013.5 hectares from the total forest area of Kalimantan 53,054,900 hectares or 0.458%. K-Means 5 clusters between Sumatra Island and Java Island are the first Cluster are 82 types of Wood, the second cluster is Acacia Wood, the third cluster is Eucalyptus Wood, the fourth cluster is Sengon Wood and the fifth cluster is Mixed Jungle Wood, Mahogany and Teak.

Keywords—fire, k-means, biodiversity, geopandas, forest

1. PENDAHULUAN

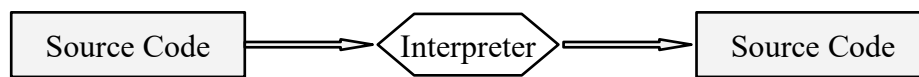
Kebakaran adalah salah satu bencana yang paling sering terjadi dan memiliki tingkat kerugian yang tidak sedikit. Indonesia termasuk sebagai salah satu negara di dunia yang memiliki kawasan hutan yang cukup luas diantaranya terdapat pada pulau Jawa, Sumatera, Kalimantan, Sulawesi dan Papua. Kerugian yang dapat ditimbulkan oleh kebakaran hutan dapat mempengaruhi kualitas udara, lahan, air, habitat flora dan fauna yang ada, dan berkurangnya produksi oksigen hingga musnahnya keanekaragaman hayati[2]. Pencegahan kebakaran sejak dini merupakan salah satu upaya yang dapat dilakukan untuk mengurangi kerusakan yang diakibatkan oleh kebakaran yang telah menyebar[3]. Kebakaran hutan dapat dipengaruhi oleh beberapa faktor diantaranya yaitu keadaan suhu, kecepatan angin, curah hujan maupun faktor dari manusia itu sendiri. Pentingnya akan kesadaran dalam menjaga hutan tentu merupakan hal yang sangat penting untuk dapat menanggulangi masalah kebakaran hutan.

Kebakaran merupakan sebuah reaksi oksidasi yang prosesnya cepat dari satu bahan dengan disertai munculnya api. Berdasarkan sumber dari NFPA (*National Fire Protection Association*) kebakaran merupakan proses oksidasi yang melibatkan oksigen, bahan bakar dan sumber panas[10]. Kebakaran hutan disebabkan oleh dua faktor yaitu faktor alam dan faktor manusia. Faktor alam yang menjadi penyebab terjadinya kebakaran hutan adalah kondisi cuaca, sedangkan faktor manusia adalah perubahan penggunaan lahan hutan maupun kebiasaan kurang baik seperti membuang korek api yang masih menyala dan membuang puntung rokok yang masih membara[9]. Dalam proses untuk pemetaan persebaran titik api pada kebakaran hutan diperlukan pengolahan data yang baik dengan analisis data mining. Data mining merupakan proses analisis terhadap data yang besar untuk dapat menemukan informasi yang berguna. sumber data dari data mining dapat berupa *database*, *warehouse*, *Web*, repositori dan informasi lain -lain. Metode data mining yang digunakan yaitu menggunakan algoritma *K-Mean Clustering*. *K-Mean Clustering* merupakan metode data clustering dan *non-hierarchical data Clustering* untuk mempartisi data kedalam bentuk satu atau lebih sehingga data yang memiliki karakteristik yang sama akan menjadi satu *cluster* dan data yang berbeda akan dikelompokkan dalam *cluster* yang lain[1].

Data mining merupakan langkah analisis terhadap sekumpulan data yang berukuran besar untuk mendapatkan hubungan antar data tersebut dan merangkumnya dalam bentuk yang mudah dipahami. Tujuan utama dari data mining adalah menggali pengetahuan dari data yang ada. Data mining adalah salah satu proses utama dari *Knowledge Discovery in Database* (KDD) proses penemuan pola-pola tertentu dari sebuah data atau basis data yang berukuran besar untuk memperoleh informasi yang sangat berguna[6]. Data mining dibagi menjadi 6, yaitu Pembentukan Pola Data, Estimasi Numerik, Prediksi, Klasifikasi, Pengklasteran dan Asosiasi[13].

Algoritma *K-Means* merupakan teknik data mining dengan menggunakan metode *Clustering* atau pengelompokan pemodelan tanpa pembelajaran dan dilakukan secara partisi. Data pada algoritma *K-Means* akan dikelompokkan menjadi beberapa bagian kelompok dan setiap kelompok memiliki ciri-ciri yang mirip satu sama lain, namun memiliki ciri-ciri yang berbeda dengan kelompok lain[2]. Algoritma dasar yang digunakan pada metode *K-Means* adalah dengan

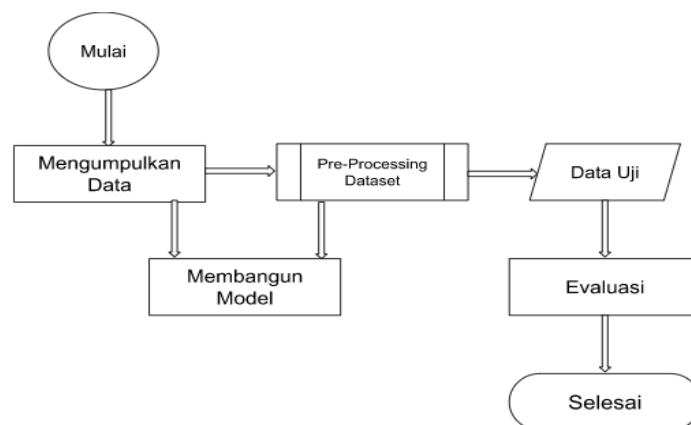
menentukan jumlah *Cluster*, pengalokasian data kedalam *cluster random*, penghitungan *centroid* pada setiap *cluster* dan dilakukan proses ujicoba kembali jika terdapat nilai yang melewati ambang batas. *Python* merupakan bahasa pemrograman yang populer digunakan oleh para programmer karena bahasa ini sangat mudah dimengerti. *Python* banyak digunakan di bidang *Data Science*, *Machine Learning*, *Artificial Intelligent*, *Face Recognition*, dan di bidang lainnya. *Python* sangat mudah dipahami karena struktur datanya yang sederhana, *python* juga merupakan pemrograman *interpreter* yang artinya program bisa langsung dieksekusi sesuai instruksi yang ditulis dalam bahasa pemrograman tanpa perlu mengubahnya menjadi kode objek seperti *Compiler*.

Gambar 1 *Interpreter*

Selain struktur data yang mudah dipahami *Python* juga banyak menyediakan modul dan library yang banyak. Modul/ Library *Python* yang populer saat ini adalah *Django*, *Web Framework*, *Scipy* dan *Scikit*, *Tornado*, *Celery*, *OpenCV Python*, *Matplotlib*, *BioPython*, *TensorFlow*, dan *Geopandas*. *Python* merupakan bahasa pemrograman yang bersifat *multi-platform*, yaitu *python* dapat berjalan pada berbagai jenis platform sistem operasi yang tersedia pada saat ini. Dengan seiring perkembangan *python* dapat di porting ke dalam mesin virtual *Java* dan *.NET*. *Porting* merupakan proses mengadaptasi perangkat lunak sehingga program dapat berjalan pada perangkat lunak yang berbeda desain dengan aslinya. Contoh *porting python* yaitu *Jython*. Pada saat ini *Python* tersedia secara bebas untuk penggunaan komersial dan nonkomersial dan berlisensi PSF v2[8].

2. METODE PENELITIAN

Penelitian ini dibagi menjadi dua bagian yaitu studi literatur dan membangun model. Studi literature bertujuan sebagai landasan teori penelitian. Adapun cakupan studi literatur adalah jurnal, artikel dan buku yang memuat tentang penerapan algoritma *Machine Learning K-Means Clustering*. Pada penelitian klasifikasi kebakaran hutan dengan menggunakan kombinasi antara algoritma *K-Nearest Neighbor* dengan *manhattan* dapat mengidentifikasi jenis titik api[5], sedangkan pada penelitian klasifikasi kebakaran menggunakan algoritma *Naive Bayes* dapat menganalisa titik panas kebakaran hutan dengan pengaruh variabel temperatur dan kondisi cuaca[12], untuk menampilkan visualisasi area yang memiliki potensi kebakaran hutan menggunakan analisis model spasial[7]. Pada studi literatur tersebut belum ada variabel tentang pola pemetaan keanekaragaman hayati hutan pada kebakaran hutan sehingga pada penelitian ini menambahkan variabel dampak kebakaran hutan pada populasi keanekaragaman hayati.



Gambar 2 Alur Proses Penelitian

Proses pengumpulan data dilakukan dengan data sekunder berupa Dataset persebaran titik panas di Indonesia dari Fire Information for Resource Management System (FIRMS) NASA berdasarkan standard processing deteksi titik panas Indonesia tahun 2021, Statistik produksi kehutanan tahun 2019 dari Badan Pusat Statistik Indonesia, Statistik lingkungan hidup Indonesia tahun 2019 hutan dan perubahan iklim dari Badan Pusat Statistik Indonesia.

2.1 Dataset

Pada pengolahan data pemetaan keanekaragaman hayati hutan diperlukan sumber data seperti dataset. Dataset merupakan kumpulan objek data yang merepresentasikan suatu topik tertentu[11]. Penelitian ini menggunakan beberapa dataset seperti Dataset persebaran titik panas di Indonesia dari Fire Information for Resource Management System (FIRMS) NASA official website untuk deteksi titik panas Indonesia tahun 2021 terdapat jumlah baris data 14.210 dengan 15 variabel data yang terdiri dari variabel latitude, longitude, brightness, scan, track, acq_date, acq_time, satellite, instrument, confidence, version, bright_t31, frp, daynight dan type.

Tabel 1 Sampel dataset persebaran titik panas di Indonesia

Latitude	Longitude	Brightness	Scan	Track	Acq_Date	Acq_Time	Satellite
-3,8988	122,4209	312,2	1,8	1,3	01/01/2021	239	Terra
-5,2174	119,4392	317,1	1,2	1,1	01/01/2021	239	Terra

Instrument	Confidence	Version	Bright_T31	Frp	Daynight	Type
MODIS	39	6,03	285,9	16	D	2
MODIS	33	6,03	294,3	7,7	D	0

Statistik produksi kehutanan tahun 2019 kategori Kayu Bulat dari Laporan Badan Pusat Statistik Indonesia dengan jumlah baris data dari FIRMS adalah 88 dengan 7 variabel data yang terdiri dari jenis_produksi_hutan, sumatera, jawa, bali_dan_nusa_tenggara, kalimantan, sulawesi, maluku_dan_papua.

Tabel 2 Sampel jenis produksi hutan tahun 2019 (dalam M³)

Jenis_Produksi_Hutan	Sumatera	Jawa	Bali_Dan_Nusa_Tenggara	Kalimantan	Sulawesi	Maluku_Dan_Papua
Akasia	28573433,84	100747,22	1885,26	4015313,83	532,47	0
Ekaliptus	6983702,23	0	2957,6	2023461,92	2699,62	0
Sengon	5189,34	5431029,46	838,34	29091,79	841,33	1726,50
Rimba Campuran	153806,69	1099215,8	34293,26	215204,04	106886,39	341478,18

Statistik lingkungan hidup Indonesia tahun 2019 hutan dan perubahan iklim dari Badan Pusat Statistik Indonesia yang terdiri dari Dataset kondisi kawasan hutan tetap (Ribuan Hektar), Dataset kondisi cuaca, Dataset kawasan hutan dan Dataset Rekapitulasi Kebakaran Hutan (Hektar). Dataset tersebut terbagi dalam area propinsi, kemudian dijadikan menjadi kategori perpulau (sesuai Tabel 3, Tabel 4, dan Tabel 5).

Tabel 3 Sampel dataset kondisi kawasan hutan (ribu Ha)

Jenis Hutan	Sumatera	Jawa	Bali_Dan_Nusa Tenggara	Kalimantan	Sulawesi	Maluku_Dan_Papua
Hutan Konservasi	5062,50	492,90	455,70	4956,40	1833,30	9087,00
Hutan Lindung	5616,90	734,90	1210,90	7031,60	4408,70	10658,00
Hutan_Produksi_Tetap	7370,90	1389,30	440,80	10794,90	1153,60	8052,90

Tabel 4 Sampel Dataset kondisi cuaca

Kondisi Cuaca	Sumatera	Jawa	Bali_Dan_Nusa Tenggara	Kalimantan	Sulawesi	Maluku_Dan_Papua
suhu udara rata rata 2018	27,64	28,05	28,30	27,96	27,55	27,50
jumlah curah hujan mm	2619,70	1803,00	1519,00	2637,80	1991,67	2366,00
kelembaban udara rata rata persen	82,20	74,00	75,67	80,40	80,50	83,50

Tabel 5 Sampel dataset rekapitulasi kebakaran hutan 2016 – 2018 (Hektar)

Kebakaran Tahun	Sumatera	Jawa	Bali_Dan_Nusa Tenggara	Kalimantan	Sulawesi	Maluku_Dan_Papua
tahun 2016	151371,55	0,00	9044,16	62898,56	19367,58	195051,34
tahun 2017	23789,93	11793,02	88727,15	18260,09	5950,55	33872,63
tahun 2018	79967,68	13322,57	95112,89	243013,50	15946,37	104112,32

2.2 Pre-processing Dataset

Tahap preprocessing data pada penelitian ini menggunakan metode pengolahan data secara manual kemudian dilanjutkan dengan menggunakan *Python* dari *Google Collabs*.

2.2.1 Metode pengolahan data secara manual

Metode pengolahan data pada tahap ini dilakukan dengan cara memilih kategori data yang akan diolah seperti pada dataset Statistik produksi kehutanan tahun 2019 dari Badan Pusat Statistik Indonesia terdapat dataset produksi kayu bulat, produksi kayu olahan, dan produksi hasil hutan non kayu menjadi satu jenis dataset saja yaitu dataset produksi kayu. Kemudian dilanjutkan dengan pengolahan data dari statistik lingkungan hidup tahun 2019 yang terdiri dari tabel data: Suhu udara, Rata-rata lama penyinaran matahari, jumlah curah hujan dan jumlah hari hujan, Kelembaban udara, Kecepatan angin, Tekanan udara di stasiun pengamatan Badan Meteorologi Klimatologi dan Geofisika (BMKG) menurut propinsi tahun 2017 sampai tahun 2018.

Hasil pengolahan awal data dari statistik lingkungan hidup tahun 2019 menjadi satu tabel sesuai tabel 4 yaitu Dataset kondisi cuaca dengan variabel sejumlah 10 yaitu propinsi, stasiun_bmkg, suhu_udara_rata-rata_2018, durasi_penyinaran_matahari_jam, jumlah_curah_hujan_mm, jumlah_hari_hujan, kelembaban_udara_rata_rata_persen, kecepatan_angin_meter/detik, tekanan_udara_milibar dan indek_kualitas_udara. Kemudian dilanjutkan pengolahan data pada dataset kawasan hutan produksi dan kawasan lain setiap propinsi di indonesia sesuai dengan Tabel 5. Dataset Kawasan Hutan (Ribuan Hektar) dengan variabel pada tahap ini adalah propinsi, hutan_konservasi, hutan_produksi_terbatas, hutan_produksi_tetap, hutan_produksi_yang_dapat_dikonversi, dan area_hutan_penggunaan_lain. Pengolahan dataset pada tahap ini juga

mengumpulkan data tentang luas kebakaran hutan antara tahun 2014 sampai dengan 2019 sesuai dengan tabel Dataset Rekapitulasi Kebakaran Hutan 2014 – 2019 (Hektar)

2.2.2 Metode pengolahan data menggunakan pemrograman dari google collabs

Metode pengolahan data pada tahap ini diawali dari pengolahan dataset produksi kehutanan hingga dataset persebaran titik panas di Indonesia dari *Fire Information for Resource Management System* (FIRMS) NASA official website dari variabel latitude dan longitude diubah menjadi variabel *geometry* menggunakan pemrograman python dengan bantuan *library geopandas, matplotlib pyplot dan pandas*. Kemudian dilanjutkan dengan pengolahan data untuk mendapatkan *Cluster* dari Jenis produksi hutan dan kondisi cuaca terhadap area kebakaran hutan setiap pulau besar di Indonesia

2.2.3 Algoritma K-Means Clustering

Algoritma *K-Means* merupakan teknik data mining yaitu metode *Clustering* atau pengelompokan yang proses pemodelannya tanpa supervisi/pembelajaran serta metode pengelompokan datanya dilakukan secara partisi. Data pada algoritma *K-Means* akan dikelompokkan menjadi beberapa bagian kelompok dan setiap kelompok memiliki ciri-ciri yang mirip satu sama lain, namun memiliki ciri-ciri yang berbeda dengan kelompok lain[2]. Kluster adalah kumpulan data yang telah disimpan dimana setiap data memiliki kemiripan satu dengan yang lainnya. Kluster memiliki memiliki keunikan yaitu tidak adanya variabel target, tidak melakukan klasifikasi, perkiraan maupun prediksi nilai dari variabel target. sehingga pengolahan data dengan pengklusteran dilakukan untuk proses pembagian terhadap keseluruhan data menjadi kelompok yang memiliki kemiripan akan bernilai maksimal, yang tidak memiliki kemiripan akan bernilai minimal dan adanya rata-rata nilai antara rata-rata kemiripan[4].

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Proses pengumpulan data dilakukan dengan cara mengambil data sekunder berupa Dataset persebaran titik panas di Indonesia dari *Fire Information for Resource Management System* (FIRMS) NASA official website berdasarkan standard processing deteksi titik panas Negara Indonesia tahun 2021. FIRMS NASA membagikan data kebakaran aktif mendekati real time setelah tiga jam pengamatan satelit dengan menggunakan *Moderate Resolution Imaging Spectroradiometer* (MODIS) NASA dan *Visible Infrared Imaging Radiometer Suite* (VIIRS). Dataset persebaran titik panas di Indonesia didownload melalui website <https://firms.modaps.eosdis.nasa.gov/country/> kemudian memilih Instrument MODIS. Selanjutnya tampil file tahun dan data semua negara. Setelah download kemudian ekstrak file dan memilih Dataset persebaran titik api di Indonesia. *File* hasil ekstrak dalam bentuk format *microsoft office excel*.

Statistik produksi kehutanan tahun 2019 dari Badan Pusat Statistik Indonesia diperoleh melalui website Badan Pusat Statistik Indonesia <https://www.bps.go.id/> kemudian memilih Menu Publikasi. Setelah memilih Menu Publikasi maka website akan menampilkan data Publikasi yang telah diterbitkan oleh Badan Pusat Statistik Indonesia. Daftar publikasi yang telah diterbitkan oleh Badan Pusat Statistik Indonesia sampai dengan tahun 2022 pada penelitian ini sebanyak 4691 *File*. Untuk memilih *file* data yang sesuai dengan penelitian ini maka perlu memasukan tahun dan kata kunci (pada penelitian ini memerlukan data dengan pilihan tahun 2019 dan kata kunci statistik kehutanan) kemudian dilanjutkan dengan *download file* Statistik Produksi Kehutanan 2019.

Statistik lingkungan hidup Indonesia tahun 2019 hutan dan perubahan iklim dari Badan Pusat Statistik Indonesia diperoleh melalui website Badan Pusat Statistik Indonesia <https://www.bps.go.id/>, kemudian memilih Menu Publikasi dan dilanjutkan mendownload *file* Statistik Lingkungan Hidup Indonesia 2019.

3.2 Dataset

Dataset persebaran titik panas di Indonesia dari *Fire Information for Resource Management System* (FIRMS) NASA official website berdasarkan standard processing deteksi titik panas Negara Indonesia tahun 2021. Dataset tersebut berupa *file microsoft office excel*, kemudian reformat ke file CSV. Statistik produksi kehutanan tahun 2019 kategori Kayu Bulat dari Badan Pusat Statistik Indonesia. Jumlah baris data dari dataset produksi kehutanan tahun 2019 kategori kayu bulat adalah 88 dengan 7 variabel data yang terdiri dari jenis_produk_hutan, sumatera, jawa, bali_dan_nusa_tenggara, kalimantan, sulawesi, maluku_dan_papua, Statistik lingkungan hidup Indonesia tahun 2019 hutan dan perubahan iklim dari Badan Pusat Statistik Indonesia yang terdiri dari Dataset kondisi kawasan hutan tetap (Ribuan Hektar), Dataset kondisi cuaca, Dataset kawasan hutan dan Dataset Rekapitulasi Kebakaran Hutan (Hektar), dan Dataset kondisi cuaca tahun 2018 dari Statistik lingkungan hidup Indonesia dalam bentuk file PDF, kemudian konversi dataset ke format excel dan langkah selanjutnya konversi dataset ke format CSV.

3.3 Algoritma K-Means Clustering

Pengolahan dataset persebaran titik panas di Indonesia dari *Fire Information for Resource Management System* (FIRMS) NASA official website berdasarkan *standard processing* deteksi titik panas Negara Indonesia tahun 2021 menggunakan pemrograman *Python Google Collabs* dengan visualisasi sebagai berikut:

```
import pandas as pd
import geopandas
import matplotlib_inline
import matplotlib.pyplot as plt
import numpy as np
world = geopandas.read_file(geopandas.datasets.get_path("naturalearth_lowres"))
world.loc[world["continent"] == "Asia"]
```

	pop_est	continent	name	iso_a3	gdp_md_est	geometry
5	18556698	Asia	Kazakhstan	KAZ	460700.00	POLYGON (((87.35997 49.21498, 86.59878 48.54918...
6	29748859	Asia	Uzbekistan	UZB	202300.00	POLYGON (((55.96819 41.30864, 55.92892 44.99586...
8	260580739	Asia	Indonesia	IDN	3028000.00	MULTIPOLYGON (((141.00021 -2.60015, 141.01706 ...
24	1291358	Asia	Timor-Leste	TLS	4975.00	POLYGON ((124.96868 -8.89279, 125.08625 -8.656...

Gambar 3 Pengolahan dataset persebaran titik api dengan *world continent asia*

Pada proses pengolahan visualisasi persebaran titik api dengan menggunakan pemrograman *python* dan *library geopandas* dimulai dari *import library* dan visualisasi dataset *world*, dilanjutkan dengan menampilkan dataset untuk Indonesia dengan *continent Asia* (sesuai Gambar 3), dilanjutkan dengan *layer* kedua berupa dataset persebaran titik api (sesuai Gambar 4)

```
[ ] import pandas as pd
data_2 = pd.read_csv('modis_2021_Indonesia.csv')
print(data_2)
```

	latitude	longitude	brightness	scan	track	acq_date	acq_time	\
0	-3.8988	122.4209	312.2	1.8	1.3	2021-01-01	239	
1	-5.2174	119.4392	317.1	1.2	1.1	2021-01-01	239	
2	-8.9146	116.7551	313.1	1.0	1.0	2021-01-01	240	
3	-1.3673	121.2896	328.5	1.0	1.0	2021-01-01	534	
4	-1.3660	121.2988	334.5	1.0	1.0	2021-01-01	534	
...	...	...	...	...	...	...	...	
14205	-7.8299	113.1501	309.1	1.3	1.1	2021-12-31	259	
14206	-4.5610	136.8871	318.2	1.2	1.1	2021-12-31	420	
14207	-5.8019	139.6112	309.6	1.0	1.0	2021-12-31	420	
14208	-4.5166	136.8480	313.7	1.2	1.1	2021-12-31	420	
14209	-4.5467	136.7750	309.2	1.2	1.1	2021-12-31	420	
...	...	...	...	...	...	...	...	
0	Terra	MODIS	39	6.03	285.9	16.0	D	
1	Terra	MODIS	33	6.03	294.3	7.7	D	
2	Terra	MODIS	0	6.03	293.8	7.0	D	
3	Aqua	MODIS	85	6.03	296.6	21.7	D	
4	Aqua	MODIS	88	6.03	296.6	29.9	D	
...	...	...	...	...	...	...	...	
14205	Terra	MODIS	60	6.03	278.2	7.4	D	

Gambar 4 Pengolahan dataset persebaran titik api dengan dataset dari FIRMS

Data yang telah ditampilkan berupa *dataframe*, maka untuk dapat memvisualisasi dalam bentuk *map*, dataset tersebut ditambahkan variabel *geometry* yang terdiri dari subvariabel *latitude* dan *longitude* (sesuai Gambar 5)

```
[ ] import geopandas as gpd
```

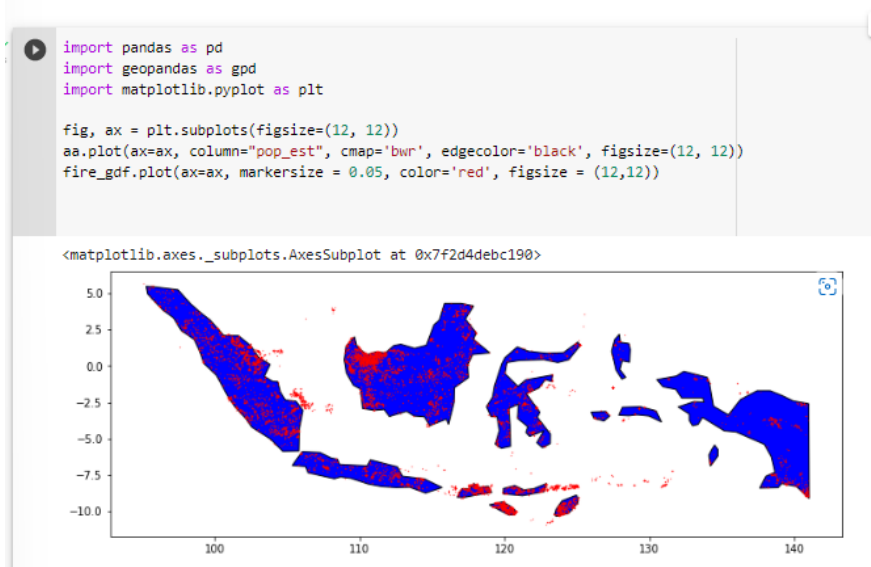
```
[ ] fire_gdf = gpd.GeoDataFrame(data_2, geometry = gpd.points_from_xy(data_2['longitude'], data_2['latitude']))
```

fire_gdf

tness	scan	track	acq_date	acq_time	satellite	instrument	confidence	version	bright_t31	frp	daynight	type	geometry
312.2	1.8	1.3	2021-01-01	239	Terra	MODIS	39	6.03	285.9	16.0	D	2	POINT (122.42090 -3.89880)
317.1	1.2	1.1	2021-01-01	239	Terra	MODIS	33	6.03	294.3	7.7	D	0	POINT (119.43920 -5.21740)
313.1	1.0	1.0	2021-01-01	240	Terra	MODIS	0	6.03	293.8	7.0	D	0	POINT (116.75510 -8.91460)
328.5	1.0	1.0	2021-01-01	534	Aqua	MODIS	85	6.03	296.6	21.7	D	0	POINT (121.28960 -1.36730)
334.5	1.0	1.0	2021-01-01	534	Aqua	MODIS	88	6.03	296.6	29.9	D	0	POINT (121.29880 -1.36600)
...	...	...	...	...	...	...	...	...	...	...	...	...	...
309.1	1.3	1.1	2021-12-31	259	Terra	MODIS	60	6.03	278.2	7.4	D	0	POINT (113.15010 -7.82990)

Gambar 5 Pengolahan dataset persebaran titik api *GeoDataFrame*

Pengolahan dataset hasil format *GeoDataFrame* kemudian ditampilkan dan ukuran sebaran titik api menjadi 5% dan memberikan label warna merah. Selanjutnya menampilkan *map* Indonesia dengan *Geopandas* dan persebaran titik api kemudian dua *layer* tersebut digabungkan menjadi satu *map* (sesuai Gambar 6)



Gambar 6 Visualisasi map Indonesia dan persebaran titik api

Pengolahan dataset Statistik produksi kehutanan tahun 2019 kategori Kayu Bulat dari Badan Pusat Statistik Indonesia menggunakan *Python Google Collabs* dengan menggunakan *library pydrive* untuk menghubungkan antara dataset yang tersimpan di *Google drive*, dan *library pandas* untuk menampilkan dataset di Python (sesuai Gambar 7).

```
[ ] df = pd.read_csv('dataset_produk_hutan.csv')
df.head(88)
```

	jenis_produk_hutan	sumatera	jawa	bali_dan_nusa_tenggara	kalimantan	sulawesi	maluku_dan_papua
0	Akasia	28573433.84	100747.22	1885.26	4015313.83	532.47	0.00
1	Ekaliptus	6983702.23	0.00	2957.60	2023461.92	2699.62	0.00
2	Sengon	5189.34	5431029.46	838.34	29091.79	841.33	1726.50
3	Meranti	104168.76	0.00	843.16	2951370.18	47856.07	421451.57
4	Rimba Campuran	153806.69	1099215.80	34293.26	215204.04	106886.39	341478.18
...	...	...	...	...	...	...	...
83	Berumbung	0.00	0.00	0.00	0.00	0.00	5.78
84	Melur	0.00	0.00	0.00	0.00	0.00	3.50
85	Pelawan	0.00	0.00	0.00	0.00	3.50	0.00
86	Durian	1.30	0.00	0.00	0.00	0.00	0.00
87	Lainnya	173189.94	323325.72	28001.63	49477.43	849.80	27189.30

88 rows x 7 columns

Gambar 7 Visualisasi dataset produksi hutan menggunakan *library geopandas*

Selanjutnya menggunakan *library sklearn* untuk *clustering K-means* dan *Pyplot* untuk menampilkan grafik pemetaan dataset di *Python*. Pada hasil pemetaan nilai *min* dan *max* dari jenis produksi hutan setiap pulau dengan menggunakan *code* :

```
pulau_sumatera = df['sumatera']
pulau_jawa = df['jawa']
pulau_bali_dan_nusa_tenggara = df['bali_dan_nusa_tenggara']
pulau_kalimantan = df['kalimantan']
pulau_sulawesi = df['sulawesi']
pulau_maluku_dan_papua = df['maluku_dan_papua']
print(pulau_sumatera, pulau_jawa, pulau_bali_dan_nusa_tenggara, pulau_kalimantan, p
      pulau_sulawesi, pulau_maluku_dan_papua)

sm_min = pulau_sumatera.min(); sm_max = pulau_sumatera.max()
jw_min = pulau_jawa.min(); jw_max = pulau_jawa.max()
bn_min = pulau_bali_dan_nusa_tenggara.min()
bn_max = pulau_bali_dan_nusa_tenggara.max()
kl_min = pulau_kalimantan.min(); kl_max = pulau_kalimantan.max()
sw_min = pulau_sulawesi.min(); sw_max = pulau_sulawesi.max()
mp_min = pulau_maluku_dan_papua.min(); mp_max = pulau_maluku_dan_papua.max()
print(sm_min, sm_max, jw_min, jw_max, bn_min, bn_max, kl_min, kl_max, sw_min, sw_ma
      x, mp_min, mp_max)

data.sort_values('sumatera', inplace=True) #'sumatera' diganti pulau lainya
produksi_hutan = data[data['sumatera'] <= 0]
produksi_hutan
```

maka ditampilkan nilai produksi minimal dan maksimal setiap pulau sesuai dengan tabel 6 berikut ini:

Tabel 6 Nilai produksi hutan minimal dan maksimal setiap pulau di Indonesia

Pulau	Produksi Minimal	Produksi Maksimal
Sumatera	0.0 (Balsa, 56 dari 88 row)	28573433.84 (Akasia)
Jawa	0.0 (Gerunggung, 67 dari 88 row)	5431029.46 (Sengon)
Bali dan Nusa Tenggara	0.0 (Gerunggung, 73 dari 88 row)	34293.26 (Rimba Campuran)
Kalimantan	0.0 (Balsa, 46 dari 88 row)	4015313.83 (Akasia)
Sulawesi	0.0 (Balsa, 56 dari 88 row)	106886.39 (Rimba Campuran)
Maluku dan Papua	0.0 (Balsa, 51 dari 88 row)	739584.8 (Merbau)

Selanjutnya untuk menghitung nilai rata-rata jenis produksi hutan di setiap pulau menggunakan *library numpy* sebagai berikut:

```
import numpy as np
mean_sm = np.mean(df['sumatera']); mean_jw = np.mean(df['jawa'])
mean_bn = np.mean(df['bali_dan_nusa_tenggara'])
mean_kl = np.mean(df['kalimantan']); mean_sw = np.mean(df['sulawesi'])
mean_mp = np.mean(df['maluku_dan_papua'])
print(mean_sm, mean_jw, mean_bn, mean_kl, mean_sw, mean_mp)
```

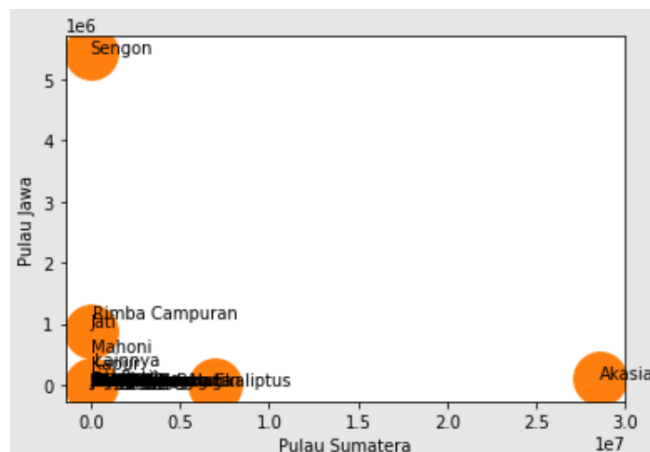
menghasilkan nilai rata-rata jenis produksi hutan setiap pulau sebagai berikut :

Rata-rata Pulau Sumatera	= 410978.73272727267
Rata-rata Pulau Jawa	= 100645.61988636364
Rata-rata Pulau Bali dan Nusa Tenggara	= 939.5342045454546
Rata-rata Pulau Kalimantan	= 120860.22613636364
Rata-rata Pulau Sulawesi	= 2215.926363636364
Rata-rata Pulau Maluku dan Papua	= 22436.109204545457

Selanjutnya menggunakan *library sklearn, numpy* dan *matplotlib* untuk menampilkan grafik pemetaan produksi hutan pulau sumatera dan pulau jawa menggunakan algoritma *K-means* dengan 5 cluster sesuai Gambar 8.

```
from sklearn.cluster import KMeans
import numpy as np
import matplotlib.pyplot as plt
sumatera_jawa = np.column_stack((pulau_sumatera, pulau_jawa))
km_res = KMeans(n_clusters=5).fit(sumatera_jawa)
cluster = km_res.cluster_centers_
plt.scatter(pulau_sumatera, pulau_jawa)
plt.scatter(cluster[:,0], cluster[:,1], s=1000)

for k, row in produksi_hutan.iterrows():
    plt.text(row['sumatera'],
            row['jawa'],
            row['jenis_produksi_hutan'])
plt.xlabel('Pulau Sumatera')
plt.ylabel('Pulau Jawa')
```



Gambar 8 Pemetaan Kmeans jenis produksi hutan di pulau sumatera dan pulau jawa

Untuk menampilkan data *clustering* (sesuai Gambar 9) digunakan kode *python*:

```
km = KMeans(n_clusters=5)
y_pengkusteran = km.fit_predict(data[['sumatera', 'jawa']])
data['cluster'] = y_pengkusteran
data.head(88)
```

index	jenis_produk_hutan	sumatera	jawa	bali_dan_nusa_tenggara	kalimantan	sulawesi	maluku_dan_papua	cluster
5	Jati	0.0	944567.84	6249.88	0.0	2554.58	0.0	4
8	Mahoni	0.0	555232.54	789.54	0.0	0.0	0.0	4
4	Rimba Campuran	153806.69	1099215.8	34293.26	215204.04	106886.39	341478.18	4
2	Sengon	5189.34	5431029.46	838.34	29091.79	841.33	1726.5	3
1	Ekaliptus	6983702.23	0.0	2957.6	2023461.92	2699.62	0.0	2
0	Akasia	28573433.84	100747.22	1885.26	4015313.83	532.47	0.0	1
43	Balsa	0.0	4652.02	0.0	0.0	0.0	0.0	0
64	Afrika	0.0	250.78	0.0	0.0	0.0	0.0	0

Gambar 9 Pengklusteran K-means jenis produksi hutan di pulau sumatera, pulau jawa

Hasilnya berupa clustering jenis produksi hutan untuk pulau sumatera dan pulau jawa adalah Klaster 1 (82 dari 88 jenis produksi hutan), Klaster 2 (Akasia), Klaster 3 (Ekaliptus), Cluster 4 (Sengon) dan Cluster 5 (Rimba Campuran, Mahoni, dan Jati). Selanjutnya menggunakan *library sklearn* untuk *clustering Kmeans* dan *Pyplot* untuk menampilkan grafik pemetaan dataset di *Python*. Pada hasil pemetaan nilai *min* dan *max* dari kondisi cuaca hutan setiap pulau dengan menggunakan *code* :

```

suhu_udara = df['suhu_udara_rata_rata_2018']
penyinaran_matahari = df['durasi_penyinaran_matahari_jam']
curah_hujan = df['jumlah_curah_hujan_mm']
jumlah_hujan = df['jumlah_hari_hujan']
kelembaban_udara = df['kelembaban_udara_rata_rata_persen']
kecepatan_angin = df['kecepatan_angin_meter_perdetik']
tekanan_udara = df['tekanan_udara_milibar']
kualitas_udara = df['indek_kualitas_udara']
kebakaran_hutan = df['kebakaran_hutan_2018']
jumlah_area_hutan = df['jumlah_area_hutan']
print(suhu_udara, penyinaran_matahari, curah_hujan, jumlah_hujan, kelembaban_udara,
      kecepatan_angin, tekanan_udara, kualitas_udara, kebakaran_hutan, jumlah_area_hutan)
su_min = suhu_udara.min(); su_max = suhu_udara.max()
pm_min = penyinaran_matahari.min(); pm_max = penyinaran_matahari.max()
ch_min = curah_hujan.min(); ch_max = curah_hujan.max()
jh_min = jumlah_hujan.min(); jh_max = jumlah_hujan.max()
ku_min = kelembaban_udara.min(); ku_max = kelembaban_udara.max()
ka_min = kecepatan_angin.min(); ka_max = kecepatan_angin.max()
tu_min = tekanan_udara.min(); tu_max = tekanan_udara.max()
kud_min = kualitas_udara.min(); kud_max = kualitas_udara.max()
kbr_min = kebakaran_hutan.min(); kbr_max = kebakaran_hutan.max()
jah_min = jumlah_area_hutan.min(); jah_max = jumlah_area_hutan.max()
print(su_min, su_max, pm_min, pm_max, ch_min, ch_max, jh_min, jh_max,
      ku_min, ku_max, ka_min, ka_max, tu_min, tu_max, kud_min, kud_max,
      kbr_min, kbr_max, jah_min, jah_max)
data.sort_values('suhu_udara_rata_rata_2018', inplace=True)
kondisi_cuaca = data[data['suhu_udara_rata_rata_2018'] >= 0]; kondisi_cuaca

```

maka ditampilkan lokasi pulau setiap kondisi cuaca sesuai dengan tabel 7 berikut ini:

Tabel 7 Kondisi cuaca minimal dan maksimal setiap pulau di Indonesia

Kondisi Cuaca	Kondisi Minimal	Kondisi Maksimal
Suhu Udara	27.5 (maluku dan papua)	28.3 (bali dan nusa tenggara)
Penyinaran Matahari	1679.15 (maluku dan papua)	2850.1 (bali dan nusa tenggara)
Curah Hujan	1519.0 (bali dan nusa tenggara)	2637.8 (kalimantan)
Jumlah Hujan	74.0 (jawa)	83.5 (maluku dan papua)
Kelembaban Udara	3.25 (maluku dan papua)	5.33 (bali dan nusa tenggara)
Kecepatan Angin	3.25 (maluku dan papua)	5.33 (bali dan nusa tenggara)
Tekanan Udara	1009.73 (maluku dan papua)	1011.05 (jawa)
Kualitas Udara	76.68 (jawa)	90.87 (sulawesi)
Kebakaran Hutan (hektar)	13322.57 (jawa)	243013.5 (kalimantan)
Area Hutan (ribu ha.)	7269.5 (bali dan nusa tenggara)	53054.9 (kalimantan)

Selanjutnya untuk menghitung nilai rata-rata kondisi cuaca menggunakan *library numpy*

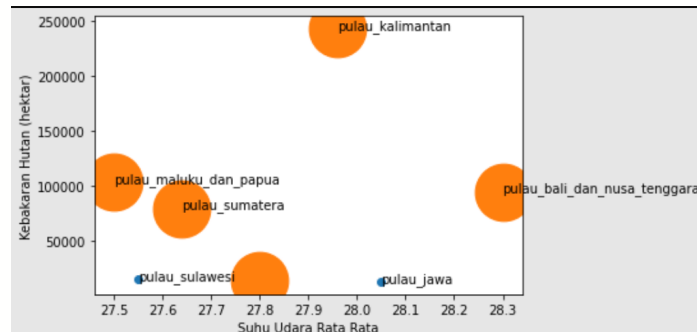
sebagai berikut:

```
import numpy as np
mean_su = np.mean(df['suhu_udara_rata_rata_2018'])
mean_pm = np.mean(df['durasi_penyinaran_matahari_jam'])
mean_ch = np.mean(df['jumlah_curah_hujan_mm'])
mean_jh = np.mean(df['jumlah_hari_hujan'])
mean_ku = np.mean(df['kelembaban_udara_rata_rata_persen'])
mean_ka = np.mean(df['kecepatan_angin_meter_perdetik'])
mean_tu = np.mean(df['tekanan_udara_milibar'])
mean_kud = np.mean(df['indek_kualitas_udara'])
mean_kbr = np.mean(df['kebakaran_hutan_2018'])
mean_jah = np.mean(df['jumlah_area_hutan'])
print(mean_su, mean_pm, mean_ch, mean_jh, mean_ku, mean_ka, mean_tu, mean_kud, mean_kbr, mean_jah)
```

menghasilkan nilai rata-rata kondisi cuaca hutan sebagai berikut :

Rata-rata Suhu Udara	=	27.833333333333332
Rata-rata Penyinaran Matahari	=	2039.6349999999995
Rata-rata Curah Hujan	=	2156.195
Rata-rata Hari Hujan	=	125.97166666666665
Rata-rata Kelembaban Udara	=	79.37833333333333
Rata-rata Kecepatan Angin	=	3.685
Rata-rata Tekanan Udara	=	1010.2183333333332
Rata-rata Kualitas Udara	=	86.61666666666667
Rata-rata Area Kebakaran Hutan	=	91912.55500000001 (hektar)
Rata-rata Jumlah Area Hutan	=	31294.100000000002 (ribu hektar)

Selanjutnya menggunakan *library sklearn, numpy* dan *matplotlib* untuk menampilkan grafik pemetaan kondisi cuaca hutan pulau sumatera dan pulau jawa menggunakan algoritma *K-Means* dengan 5 *cluster* (sesuai Gambar 10)



Gambar 10 Kmeans pemetaan kondisi cuaca hutan suhu udara dan kebakaran hutan

Untuk menampilkan data *clustering* (sesuai Gambar 11) digunakan kode *python*:

```
km = KMeans(n_clusters=5)
y_kluster = km.fit_predict(data[['suhu_udara_rata_rata_2018', 'kebakaran_hutan_2018']])
data['cluster'] = y_kluster
data.head(6)
```

	propinsi	suhu_udara_rata_rata_2018	kebakaran_hutan_2018	jumlah_area_hutan	cluster
0	pulau_sumatera	27.64	79967.68	47205.2	3
1	pulau_jawa	28.05	13322.57	13316.9	2
2	pulau_bali_dan_nusa_tenggara	28.30	95112.89	7269.5	0
3	pulau_kalimantan	27.96	243013.50	53054.9	1
4	pulau_sulawesi	27.55	15946.37	18463.6	2
5	pulau_maluku_dan_papua	27.50	104112.32	48454.5	4

Gambar 11 Pengklusteran kondisi cuaca suhu udara dan area kebakaran hutan

Hasil Klaster kondisi cuaca hutan untuk suhu udara dan area kebakaran hutan adalah Klaster 1 (Pulau Bali dan Nusa Tenggara), Klaster 2 (Pulau Kalimantan), Klaster 3 (Pulau Jawa dan Pulau Sulawesi), Klaster 4 (Pulau Sumatera), Klaster 5 (Pulau Maluku dan Papua)

4. KESIMPULAN

Hasil pemetaan *Biodiversity* jenis produksi hutan kategori kayu bulat pulau-pulau besar di Indonesia dengan menggunakan *Clustering K-means* didapatkan hasil sebagai berikut:

Produksi Hutan

Nilai produksi tertinggi Pulau Sumatera 28573433.84 m³ jenis kayu Akasia, Pulau Jawa produksi hutan tertingginya 5431029.46 m³ jenis kayu Sengon, Pulau Bali dan Nusa Tenggara produksi hutan tertingginya 34293.26 m³ jenis kayu Rimba Campuran, Pulau Kalimantan produksi hutan tertingginya 4015313.83 m³ jenis kayu Akasia, Pulau Sulawesi produksi hutan tertingginya 106886.39 m³ jenis kayu Rimba Campuran, Pulau Maluku dan Papua produksi hutan tertingginya 739584.8 m³ jenis kayu Merbau

Kondisi Cuaca

Kondisi cuaca hutan kategori suhu tertinggi rata-rata tahun adalah 28.3 derajat Celcius di Pulau Bali dan Nusa Tenggara, Durasi penyinaran matahari terbanyak 2850.1 di Pulau Bali dan Nusa Tenggara, Curah hujan tertinggi 2637.8 di Pulau Kalimantan, Jumlah hari hujan tertinggi 158.2 hari di Pulau Kalimantan, Kelembaban udara tertinggi 83.5 di Pulau Maluku dan Papua, Kecepatan Angin tertinggi 5.33 di Pulau Bali dan Nusa Tenggara, Tekanan Udara tertinggi 1011.05 di Pulau Jawa, Kualitas Udara tertinggi 90.87 di Pulau Sulawesi, total area kebakaran hutan tertinggi 243013.5 hektar di Pulau Kalimantan dan Jumlah Area Hutan tertinggi 53054.9 Ribu Hektar di Pulau Kalimantan.

Klaster Jenis Produksi Hutan

Klaster jenis produksi hutan untuk pulau sumatera dan pulau jawa dengan 5 klaster adalah Klaster 1 (82 dari 88 jenis produksi hutan), Klaster 2 (Akasia), Klaster 3 (Ekaliptus), Cluster 4 (Sengon) dan Cluster 5 (Rimba Campuran, Mahoni, dan Jati)

Klaxster Kondisi Cuaca Hutan

Klaster kondisi cuaca hutan untuk suhu udara dan area kebakaran hutan untuk 5 klaster adalah Klaster 1 (Pulau Bali dan Nusa Tenggara), Klaster 2 (Pulau Kalimantan), Klaster 3 (Pulau Jawa dan Pulau Sulawesi), Klaster 4 (Pulau Sumatera), Klaster 5 (Pulau Maluku dan Papua)

5. SARAN

Penelitian ini menggunakan sampel pada kawasan lahan hutan dan belum memetakan kategori kawasan lahan penggunaan lainya seperti lahan gambut. Untuk jenis *Biodiversity* hutan juga hanya mengambil sampel Flora jenis produksi hutan berupa jenis Pohon Kayu.

DAFTAR PUSTAKA

- [1] Agusta, Y. (2007). K-Means - Penerapan, Permasalahan dan Metode Terkait, *Jurnal Sistem dan Informatika*.
- [2] Dhewayani, F. N., Amelia, D., Alifah, D. N., Sari, B. N., & Jajuli, M. (2022). Implementasi Fakultas Ilmu Komputer | Universitas Klabat | CORIS | ISSN: 2541-2221 | E-ISSN: 2477-8079

- K-Means Clustering untuk pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM. *Jurnal Teknologi Dan Informasi*, 12. https://doi.org/Jurnal_Teknologi_dan_Informas_i
- [3] Febrianti, A. (2018). Penerapan Metode K-Means Clustering dan Support Vector Machine (SVM) dalam identifikasi Api Pada Citra Warna Digital.
- [4] Kusrini, & Luthfi, E. T. (2009). Algoritma Data Mining (T. A. Prabawati (ed.)).
- [5] Muslim, I., Karo, K., Khosuri, A., Steiven, J., Septory, I., & Supandi, P. (2022). Pengaruh Metode Pengukuran Jarak pada Algoritma k-NN untuk Klasifikasi Kebakaran Hutan dan Lahan. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 2, 1174–1182. <https://doi.org/10.30865/MIB.V6I2.3967>
- [6] Prasetyowati, E. (2017). Data Mining Pengelompokan Data untuk Informasi dan Evaluasi. In *Duta Media*. Duta Media Publishing. <https://doi.org/IKAPI> : 180/JTI/2017
- [7] Purnawan, B. I., Hasyim, A. W., Yudono, D. A., Di, S., Bobby, J., & Purnawan, I. (2021). Strategi Pengelolaan Kawasan Berpotensi Kebakaran Hutan Berbasis Pemodelan Spasial di Kabupaten Muaro Jambi. *Jurnal Sosial Dan Teknologi (SOSTECH)*, 1, 963–981. <https://doi.org/10.36418/JURNALSOSTECH.V1I9.193>
- [8] Rangkuti, Y. M., Idrus, S. I., & Tarigan, D. D. (2021). Pengantar Pemrograman Python. *Media Sains Indonesia*.
- [9] Ruhimat, M., Supriatna, N., & Kosim. (2017). Ilmu pengetahuan Sosial (Geografi, Sejarah, Sosiologi, Ekonomi). *Grafindo Media Pratama*.
- [10] Sanjoto, S. . A. F., & Komari, ana. (2019). Analisa Upaya Pencegahan dan Penanggulangan Potensi Bahaya.
- [11] Suhartini, S., Kerta Wijaya, L., & Arini Pratiwi, N. (2020). Penerapan Algoritma K-Means Untuk Pendataan Obat Berdasarkan Laporan Bulanan Pada Dinas Kesehatan Kabupaten Lombok Timur. *Infotek : Jurnal Informatika Dan Teknologi*, 2, 147–156. <https://doi.org/10.29408/JIT.V3I2.2315>
- [12] Trya Ayu Pratiwi, K. di, I., M., K., R., Agustian, S., & Sukma Negara, B. (2021). Klasifikasi Kebakaran Hutan Dan Lahan Menggunakan Algoritma Naïve Bayes Di Kabupaten Pelalawan. *Journal of Computer Engineering, System and Science*, 139–148. <https://doi.org/10.24114/CESS.V6I1.22555>
- [13] Wahyudi, M., Mashita, Saragih, R., & Solikhun. (n.d.). Data Mining : Penerapan Algoritma K-Means Clustering dan K-Medoids Cluster. *Yayasan Kita Menulis*.