

Implementation of K-Means Algorithm in Determining Clustering of Lion Air Pilots

Ahmad Jurnaidi Wahidin

Program Studi Teknologi Informasi, Universitas Bina Sarana Informatika

Email: ahmad.ajn@bsi.ac.id

Abstract

Covid-19 was declared a pandemic by WHO on March 11, 2020, and President Joko Widodo through the Head of BNPB declared the COVID-19 outbreak a National Disaster. To limit people's mobility, the government issued a Regulation of the Minister of Transportation concerning Transportation Control in which operators of transportation facilities are obliged to supervise and ensure the implementation of physical distancing. One of the sectors affected by the COVID-19 pandemic is transportation, the government is looking for a solution to minimize this impact by providing incentive stimulus to airlines and also reducing the Air Passenger Service tariff by 20% to reducing aviation fuel prices by 10%, but the government's efforts have not succeeded. . The implementation of PPKM together with the Circular Letter of the Ministry of Transportation Number SE 88 of 2021 has reduced the interest of airplane passengers. This situation made the Lion Air airline reduce 8000 employees to pilots. However, pilots are applied to fly alternately by dividing into several groups. By using the k-means algorithm which is one of the clustering methods in data mining, the division of groups can be done by partitioning the data into several clusters. so that data with the same characteristics will be grouped into one cluster. The results of processing 730 lion air pilot data with three clusters resulted in 431 cluster 1 data, 200 data cluster 2 and 99 data cluster 3, from the three clusters dominated by the age of 20-35 years with a working period of 6-11 years but with an average different average flying hours and working hours.

Keywords: *K-Means Algorithms, Data Mining, Pilot*

1. INTRODUCTION

*Coronavirus Disease 2019 (COVID-19) adalah jenis penyakit baru yang disebabkan oleh virus dari golongan coronavirus, yaitu SARS-CoV-2 yang juga sering disebut virus Corona. Para ilmuwan pertama kali mendengar tentang Sars-CoV-2 setelah virus itu terdeteksi di China. Namun, virus itu mungkin sudah menyebar lama sebelum dilaporkan, ungkap sebuah studi baru. Ilmuwan dari Pusat Pengendalian dan Pencegahan Penyakit (CDC) AS menerbitkan temuan mereka dalam jurnal *Clinical Infectious Diseases*. Secara resmi, virus itu dianggap mulai menyebar pada 31 Desember 2019, ketika otoritas kesehatan di kota Wuhan di China mengeluarkan peringatan tentang serangkaian kasus yang terkait dengan virus yang menyerang pernapasan secara misterius[1]. Kemudian ditetapkan menjadi pandemi oleh Organisasi Kesehatan Dunia (WHO) secara resmi pada 11 maret 2020 [2].*

Pemerintah Indonesia telah menetapkan Wabah Corona Virus atau Covid-19 sebagai Bencana Nasional. Status tersebut diumumkan pada 14 maret 2020 oleh Presiden melalui Kepala

Badan Nasional Penanggulangan Bencana Doni Monardo di Gedung BNPB [3]. Untuk mengurangi penyebaran virus corona, pemerintah Indonesia membuat kebijakan yang bertujuan untuk membatasi pergerakan masyarakat, salah satu diantaranya adalah kebijakan persyaratan pada penumpang dalam melakukan perjalanan menggunakan transportasi udara. PP Nomor 18 tahun 2020 tentang Peraturan Menteri Perhubungan tentang Pengendalian Transportasi dalam Rangka Pencegahan Penyebaran *Corona Virus Disease* 2019 (Covid-19) yang dikeluarkan 9 april 2021 dimana operator sarana transportasi wajib mengawasi dan memastikan penerapan jaga jarak fisik (*physical distancing*) [4].

Pandemi covid-19 membawa dampak ke berbagai sektor, termasuk perekonomian dan bisnis. Beberapa bidang yang terdampak negatif adalah perdagangan, investasi, pariwisata sampai transportasi. Isu negatif virus corona membuat banyak negara, termasuk Indonesia mencari solusi agar memperkecil dampak pelemahan ekonomi seperti penurunan kontribusi dari sektor pariwisata, beberapa cara yang dilakukan Pemerintah adalah berkoordinasi memberikan stimulus insentif kepada airlines, travel agent dan tourism representatives di luar negeri, selain kebijakan insentif transportasi, PT Angkasa Pura juga memberikan pengurangan tarif Pelayanan Jasa Penumpang Pesawat Udara (PJP2U) sebesar 20%, Pertamina juga memberikan penurunan harga avtur pada 9 destinasi sebesar 10% [5].

Dengan berbagai kebijakan tersebut, diharapkan harga tiket semakin murah dan jumlah penumpang akan meningkat sebanyak 25%. Namun, pemberlakuan PPKM berbarengan dengan keluarnya Surat Edaran Kementerian Perhubungan Nomor SE 88 Tahun 2021 yang menentukan persyaratan kesehatan untuk penerbangan dari atau ke bandara di Pulau Jawa dan Pulau Bali dengan menunjukan kartu vaksin (minimal vaksinasi dosis pertama) dan surat keterangan hasil negatif tes RT-PCR yang sampelnya diambil dalam kurun waktu 2x24 jam sebelum keberangkatan membuat penumpang pesawat berkurang.

Keadaan tersebut membuat turunnya peminat penumpang pesawat udara dan membuat maskapai penerbangan menjadi sepi penumpang, dengan sepi penumpang banyak maskapai yang akhirnya mengurangi jumlah penerbangan, mengembalikan pesawat ke *lessor* dan pengurangan jumlah karyawan. Maskapai Lion Air merupakan salah satu maskapai yang terdampak, dan manajemen memutuskan untuk melakukan pengurangan pada Agustus 2021. Dengan tidak membaiknya kondisi pandemi di Indonesia ditambah kebijakan pemerintah yang membuat semakin menurun penumpang pesawat, maskapai lion air melakukan pengurangan karyawan besar-besaran pada Agustus 2021, lion air merumahkan sekitar 8000 karyawannya.

Pengurangan karyawan berada di kisaran 5.750-8.050 dari total 23 ribu karyawan, langkah itu diambil karena pandemi COVID-19 menghantam industri penerbangan sejak awal kemunculannya. Industri penerbangan melakukan efisiensi di berbagai lini agar dapat tetap

beroperasi terbatas. "Dalam jangka waktu yang diperlukan, Lion Air Group mengumumkan pengurangan tenaga kerja dengan merumahkan karyawan (status tidak Pemutusan Hubungan Kerja/PHK) menurut beban kerja (load) di unit masing-masing yaitu kurang lebih prosentase 25-35% karyawan dari 23.000 karyawan," kata Corporate Communications Strategic of Lion Air Group, Danang Mandala Prihantoro dalam keterangannya [6]. Selain staf penerbangan, Pilot dan Pramugari juga tidak luput dari target pengurangan, namun untuk pilot konsep yang diterapkan adalah terbang secara bergantian dengan membagi beberapa kelompok terbang.

Data mining dapat diterapkan untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual [7]. *Clustering* merupakan salah satu teknik *data mining*, dan *k-means* sebagai salah satu metode *Clustering* non-hirarki dapat mempartisi data yang ada ke dalam satu atau lebih *cluster*, sehingga data yang memiliki karakteristik yang sama dikelompokkan kedalam satu *cluster*. Dengan algoritma *k-means* diharapkan dapat menyelesaikan masalah pengelompokan pada pilot lion air.

2. METHODS

Berdasarkan pada permasalahan yang diteliti, metode yang digunakan dalam penelitian ini adalah metode deskriptif dengan pendekatan kuantitatif, pemilihan deskriptif dimana penelitian deskriptif adalah penelitian yang menggunakan observasi, wawancara atau angket mengenai keadaan sekarang ini, mengenai subjek yang sedang kita teliti [8]. Sedangkan pendekatan kuantitatif dipilih untuk pendekatan penelitian yang banyak menggunakan angka-angka, mulai dari mengumpulkan data, penafsiran terhadap data yang diperoleh, serta pemaparan hasilnya [9]. Implementasi metode deskriptif kuantitatif ini bertujuan mendeskripsikan atau menjelaskan peristiwa atau kejadian yang terjadi pada saat sekarang dalam bentuk angka-angka yang bermakna.

Jenis data yang digunakan dalam penulisan ini adalah data primer, data primer adalah data yang dikumpulkan melalui pihak pertama, biasanya dapat melalui wawancara, jejak dan lain-lain [10]. Data yang didapat langsung dari divisi *Flight Operation Support* (FOS).

Teknik pengumpulan data menggunakan observasi, wawancara dan studi Pustaka. Penelitian ini merupakan penelitian yang bertujuan menjelaskan fenomena yang ada dengan menggunakan angka-angka, dimana untuk melakukan pengelompokan terhadap pilot lion air menggunakan data-data yang diolah menggunakan algoritma *k-means*.

Langkah-langkah algoritma *K-Means* adalah sebagai berikut[7]:

- a. Pilih secara acak k buah data sebagai pusat *cluster*.

- b. Jarak antara data dan pusat *cluster* dihitung menggunakan *Euclidian Distance*. Untuk menghitung jarak semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak *Euclidean* yang dirumuskan pada persamaan (1):

$$D(i,j) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + \dots + (Xki - Xkj)^2} \tag{1}$$

dimana:

$D(i,j)$ = Jarak data ke i ke pusat *cluster* j

Xki = Data ke i pada atribut data ke k

Xkj = Titik pusat ke j pada atribut ke k

- c. Data ditempatkan dalam *cluster* yang terdekat, dihitung dari tengah *cluster*.
 d. Pusat *cluster* baru akan ditentukan bila semua data telah ditetapkan dalam *cluster* terdekat.
 e. Proses penentuan pusat *cluster* dan penempatan data dalam *cluster* diulangi sampai nilai *centroid* tidak berubah lagi.

3. RESULTS AND DISCUSSION

Penelitian ini memiliki tujuan menggunakan algoritma *k-means* untuk mengelompokan data Pilot pada maskapai Lion Air, penelitian ini dibuat menggunakan data primer dimana respondennya adalah pilot lion air dengan menggunakan sebanyak 730 data yang diambil pada bulan september 2021. Adapun pengolahan data menggunakan langkah-langkah sebagai berikut:

3.1 Sumber Data

Pada penelitian ini, data bersumber dari PT. Lion Group dimana data tersebut adalah data Pilot yang masih terdaftar pada bulan 1 september 2021, data pilot ditampilkan pada table 1 dibawah ini:

Tabel 1 Data Pilot

Staff Number	131328	Crew Name	Ligar Lughina	Fleet	737
Base	CGK	Rank	CPT	Function	
Emp date	09APR13	Gender	M	Seniority No	
Birth date	21JUN90	Citizen	ID	Phone	082130333407 (Mobile)
Flight hours	225:09	Working hours	570:08	Address	Perum Nirwana Valley Residence Bogor

Data pilot yang ditampilkan pada tabel 1 tidak dapat langsung diproses, dikarenakan masih terdapat data yang tidak digunakan sehingga harus dilakukan pembersihan data (*data cleaning*), seleksi data (*data selection*) dan inisialisasi data, baru data dapat dilakukan proses *minning*.

3.2 Pengolahan Data

Tahap awal pengolahan data adalah pembersihan data, pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan [11], kemudian dilakukan seleksi data, dari data yang diperoleh penulis menghilangkan beberapa data seperti *Base, Rank, Gender, Citizen, Fleet, Function, Seniority No, Phone* dan *Address*. Dari proses diatas didapat contoh data sesuai ditunjukkan pada tabel 2.

Tabel 2 Hasil Pembersihan Data

ID	Emp Date	Birth Date	Flight Hours	Working Hours
131328	09Apr13	21Jun1990	225:09	570:08
131332	09Apr13	25Mar1984	234:12	600:10
131335	09Apr13	24Dec1989	239:50	649:21
131780	07Dec13	28Jul1993	190:16	476:04
131976	29Apr13	22Feb1993	27:20	228:42
131992	20May13	02May1968	258:39	618:43
132004	19May13	03Feb1990	35:45	324:38
132945	23May13	08Nov1958	179:04	484:30
132948	04Jul13	02Apr1988	70:14	383:39
132949	04Jul13	26Feb1992	80:21	405:42

Selanjutnya adalah transformasi data (*data transformation*), data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining* [7]. Agar data di atas dapat diolah dengan menggunakan metode *k-means Clustering*, maka data yang berjenis data *nominal* seperti *emp date* dan *birth date* harus diinisialisasikan terlebih dahulu dalam bentuk angka. Dan merubah kriteria *emp date* menjadi lama kerja serta *birth date* menjadi usia.

Tabel 3 Data Hasil Inisialisasi

Nomor Pegawai	Lama Kerja	Usia	Jam Terbang	Jam Kerja
131328	8,23	31,05	225:09	570:08
131332	8,23	37,29	234:12	600:10
131335	8,23	31,54	239:50	649:21
131780	8,18	28,37	190:16	476:04
131976	8,12	53,20	27:20	228:42
131992	8,12	31,43	258:39	618:43
132004	8,11	62,69	35:45	324:38
132945	8,00	33,27	179:04	484:30
132948	8,00	29,36	70:14	383:39
132949	8,06	30,49	80:21	405:42

Kemudian dari hasil inisialisasi data dibulatkan, dan merubah data pada kriteria jam

terbang dan jam kerja dari tipe data jam menjadi bilangan bulat.

Tabel 4 Data Inisialisasi Dengan Pembulatan Angka

Nomor Pegawai	Lama Kerja	Usia	Jam Terbang	Jam Kerja
131328	8	31	225	570
131332	8	37	234	600
131335	8	32	240	649
131780	8	28	190	476
131976	8	53	27	229
131992	8	31	259	619
132004	8	63	36	325
132945	8	33	179	484
132948	8	29	70	384
132949	8	30	80	406

Untuk mempermudah proses mining, dilakukan inisialisasi terhadap id data dan kriteria data dengan merubah data sesuai yang ditunjukkan pada table 5.

Tabel 5 Inisialisasi Data ID dan Kriteria

Nomor Pegawai	Alternatif
Lama Kerja	C1
Usia	C2
Jam Terbang	C3
Jam Kerja	C4

Setelah data ditransformasi ke dalam bentuk angka dan sudah siap untuk dilakukan proses *mining*, kemudian menentukan jumlah *cluster* yang akan digunakan, dimana pada penelitian ini penulis akan membagi data-data yang diolah kedalam tiga *cluster*.

Menentukan titik pusat (*centroid*) awal dari setiap *cluster*, pada penelitian ini titik pusat (*centroid*) awal ditentukan secara acak (*random*) dan didapat titik pusat dari awal setiap *cluster* dapat dilihat pada tabel 6.

Tabel 6 Titik Pusat (*Centroid*) Awal Setiap *Cluster*

<i>Cluster</i>	Alternatif	C1	C2	C3	C4
<i>Cluster 1</i>	131780	8	28	190	476
<i>Cluster 2</i>	52116047	10	58	168	450
<i>Cluster 3</i>	181257	4	30	218	905

Penelitian ini menggunakan metode *hard k-means* untuk mengalokasikan setiap data ke dalam suatu *cluster*, sehingga data akan dimasukkan dalam suatu *cluster* yang memiliki jarak

paling dekat dengan titik pusat dari setiap *cluster*. Untuk mengetahui *cluster* mana yang paling dekat dengan data, maka perlu dihitung jarak setiap data dengan titik pusat setiap *cluster*.

Sebagai contoh akan dihitung data baris ke-1 dengan nomoer pegawai 131328 dihitung dengan titik pusat ke-1 sebagai D(1,1), dengan titik pusat ke-2 sebagai D(1,2) dan titik pusat ke-3 sebagai D(1,3). Jarak dari data pilot pertama ke pusat *cluster* pertama dihitung sebagai berikut:

$$D(1,1) = \sqrt{(8 - 8)^2 + (31 - 28)^2 + (225 - 190)^2 + (570 - 476)^2} = 100,34939$$

Dari perhitungan diatas maka menghasilkan 100,34939 untuk jarak antara data pilot pertama dengan pusat *cluster* pertama, dan untuk mengetahui jarak mahasiswa pertama ke *cluster* kedua:

$$D(1,2) = \sqrt{(8 - 10)^2 + (31 - 58)^2 + (225 - 168)^2 + (570 - 450)^2} = 135,58023$$

Dari perhitungan diatas maka menghasilkan 135,58023 untuk jarak antara data pilot pertama dengan pusat *cluster* kedua, dan untuk mengetahui jarak mahasiswa pertama ke *cluster* ketiga:

$$D(1,3) = \sqrt{(8 - 4)^2 + (31 - 30)^2 + (225 - 218)^2 + (570 - 905)^2} = 335,09849$$

Dari perhitungan diatas maka menghasilkan 335,09849 untuk jarak antara data pilot pertama dengan pusat *cluster* ketiga.

Berdasarkan hasil dari ketiga perhitungan pilot pertama diatas maka dapat disimpulkan bahwa data pilot pertama memiliki jarak paling dekat dengan *cluster* 1, sehingga data pilot pertama pada hitungan iterasi pertama dimasukan ke dalam *cluster* 1. Hasil perhitungan untuk 10 pilot pertama ditampilkan pada tabel 7.

Tabel 7 Contoh Hasil Perhitungan Setiap Data ke Setiap *Cluster*

Alternatif	C1	C2	C3	C4	Jarak ke			Jarak terdekat ke <i>cluster</i>
					<i>Cluster</i> 1	<i>Cluster</i> 2	<i>Cluster</i> 3	
131328	8	31	225	570	100,349	135,580	335,098	1
131332	8	37	234	600	131,883	165,230	305,526	1
131335	8	32	240	649	180,125	213,225	256,982	1
131780	8	28	190	476	0,000	45,431	429,936	1
131976	8	53	27	229	296,990	262,204	702,853	2
131992	8	31	259	619	158,805	193,843	288,953	1
132004	8	63	36	325	218,499	181,874	608,793	2
132945	8	33	179	484	14,491	43,658	422,832	1
132948	8	29	70	384	151,212	121,676	541,629	2
132949	8	30	80	406	130,399	102,313	517,746	2

Setelah 730 data pilot ditempat ke *cluster* yang terdekat, selanjutnya menghitung kembali titik pusat (*centroid*) *cluster* yang baru berdasarkan rata-rata anggota yang ada pada *cluster* tersebut.

Setelah didapatkan titik pusat (*centroid*) yang baru dari setiap *cluster*, kemudian lakukan kembali dari langkah ketiga hingga titik pusat (*centroid*) dari setiap *cluster* tidak berubah lagi dan tidak ada lagi data yang berpindah dari satu *cluster* ke *cluster* yang lain.

Tabel 8 Titik Pusat (*Centroid*) Setiap *Cluster* pada Iterasi Terakhir

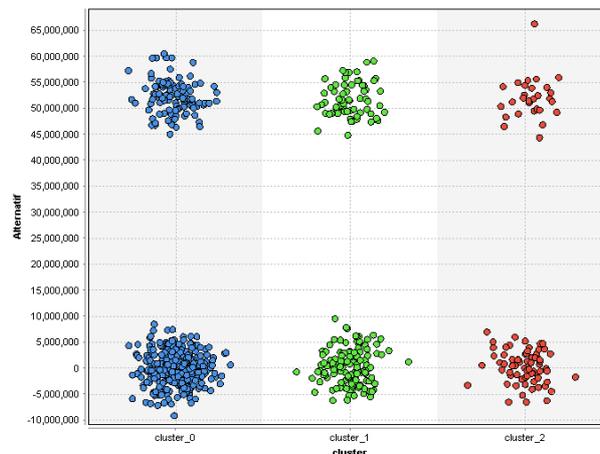
<i>Cluster</i>	C1	C2	C3	C4
<i>Cluster 1</i>	7.174	34.077	219.601	595.299
<i>Cluster 2</i>	7.525	37.275	88.945	341.875
<i>Cluster 3</i>	6.616	33.727	61.848	1117

Dari titik pusat data diatas, kembali mengalokasikan setiap data ke dalam suatu *cluster*, sehingga data akan dimasukan dalam suatu *cluster* yang memiliki jarak paling dekat dengan titik pusat dari setiap *cluster*, dan menghasilkan data yang sudah memiliki *cluster* dan contoh 10 data pilot sebagai berikut:

Tabel 9 Contoh Hasil Pengelompokan Setiap Data

Alternatif	C1	C2	C3	C4	<i>Cluster</i>
131328	8	31	225	570	<i>Cluster 1</i>
131332	8	37	234	600	<i>Cluster 1</i>
131335	8	32	240	649	<i>Cluster 1</i>
131780	8	28	190	476	<i>Cluster 1</i>
131976	8	53	27	229	<i>Cluster 2</i>
131992	8	31	259	619	<i>Cluster 1</i>
132004	8	63	36	325	<i>Cluster 2</i>
132945	8	33	179	484	<i>Cluster 1</i>
132948	8	29	70	384	<i>Cluster 2</i>
132949	8	30	80	406	<i>Cluster 2</i>

Dari tabel hasil pengelompokan setiap data diatas, maka menghasilkan sebaran data yang ditampilkan pada peta sebaran data dibawah ini:



Gambar 1 Peta Sebaran *Cluster*

Dari hasil pengolahan data dengan data mining dimana 730 data pilot lion air yang

dikelompokkan kedalam tiga *cluster*, maka dapat dianalisa sebagai berikut:

Tabel 10 Hasil Analisa *Clustering*

<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>
<i>Cluster 1</i> terdiri dari 431 Pilot, dengan pembagian <i>rank</i> : Cpt = 218 orang Fo = 213 orang	<i>Cluster 2</i> terdiri dari 200 Pilot, dengan pembagian <i>rank</i> : Cpt = 105 orang Fo = 95 orang	<i>Cluster 3</i> terdiri dari 99 Pilot, dengan pembagian <i>rank</i> : Cpt = 39 orang Fo = 60 orang
Dan berdasarkan usia: 20-35 tahun= 296 orang 36-51 tahun=103 orang 52-67 tahun= 32 orang	Dan berdasarkan usia: 20-35 tahun=124 orang 36-51 tahun=38 orang 52-67 tahun=38 orang	Dan berdasarkan usia: 20-35 tahun=69 orang 36-51 tahun=25 orang 52-67 tahun=5 orang
Dengan masa kerja: 0-5 tahun=148 orang 6-11 tahun=232 orang 12-17 tahun=48 orang 18-23 tahun=3 orang	Dengan masa kerja: 0-5 tahun=65 orang 6-11 tahun=102 orang 12-17 tahun=28 orang 18-23 tahun=5 orang	Dengan masa kerja: 0-5 tahun= 41 orang 6-11 tahun= 43 orang 12-17 tahun= 13 orang 18-23 tahun= 2 orang
Dengan rata-rata jam terbang 219,601jam, dan rata-rata jam kerja 595,299 jam	Dengan rata-rata jam terbang 88,945 jam,dan rata-rata jam kerja 341,875 jam	Dengan rata-rata jam terbang 61,8485 jam, dan rata-rata jam kerja 1117 jam

Dari hasil analisa diatas karakteristik pilot pada *cluster 1* didominasi oleh pilot dengan usia 20-35 tahun, dan untuk masa kerja paling dominan adalah 6-10 tahun. Dengan rata-rata jam terbang 219,601 jam dan jam kerja 595,299 jam.

Sedangkan *cluster 2* dan *cluster 3* untuk usia juga didominasi diusia 20-35, sampai dengan masa kerja juga sama di 6-11 tahun masa kerja. Namun untuk *cluster 2* rata-rata jam terbang 88,945 jam dan *cluster 3* 61,8485 jam dan untuk rata-rata jam kerja pada *cluster 2* adalah 341,875 sedangkan *cluster 3* adalah 1117 jam.

4. CONCLUSION

Dari penelitian yang dilakukan, algoritma *k-means* dapat mengelompokkan data kedalam suatu *cluster* berdasarkan karakteristik, sehingga data dengan karakteristik yang sama akan dikelompokkan kedalam satu *cluster*, sedangkan data dengan karakteristik yang berbeda akan dimasukan kedalam *cluster* lainnya. Hasil dari pengolahan 730 data pilot lion air dengan 3 *cluster* yang sudah ditentukan menghasilkan *cluster 1* sebanyak 431 data, *cluster 2* sebanyak 200 data dan *cluster 3* sebanyak 99 data, dari ketiga *cluster* didominasi usia 20-35 tahun dengan masa kerja 6-11 tahun namun dengan rata-rata jam terbang dan rata-rata jam kerja yang berbeda.

REFERENCES

- [1] BBC Indonesia, “Covid-19: Virus corona diduga sudah menyebar sebelum pertama kali dilaporkan di China pada Desember 2019,” *BBC.com*, Dec. 05, 2020.
- [2] I. CNBC, “WHO Nyatakan Wabah COVID-19 jadi Pandemi, Apa Maksudnya?,” *CNBCIndonesia.com*, Jakarta, Mar. 12, 2020.
- [3] Kemenkes, “Status Wabah Corona di Indonesia Ditetapkan sebagai Bencana Nasional,” *kemkes.go.id*, Jakarta, Mar. 15, 2020.
- [4] Kemenhub, “PM 18 Tahun 2020 tentang Pengendalian Transportasi dalam Rangka Pencegahan Penyebaran,” *Badan Pemeriksa Keuangan Republik Indonesia*, 2020. <https://peraturan.bpk.go.id/Home/Details/135886/permenhub-no-18-tahun-2020>.
- [5] Kemenkeu, “Ini Rencana Tambahan Insentif Pariwisata untuk Maskapai, Agen perjalanan dan Wisata Domestik,” *kemenkeu.go.id*, Jakarta, Feb. 26, 2020.
- [6] T. Hamdani, “8.000 Karyawan Lion Air Dirumahkan!,” *finance.detik.com*, Aug. 01, 2021.
- [7] A. N. Khormarudin, “Teknik Data Mining: Algoritma K-Means Clustering,” *J. Ilmu Komput.*, 2016, pp. 1–12.
- [8] E. T. Ruseffendi, “Fundamentals of Educational Research & Other Non-exact Fields,” *Bandung: Tarsito*, 2010.
- [9] S. Arikunto, “Metode penelitian kualitatif,” *Jakarta Bumi Aksara*, 2006.
- [10] S. Suharsimi Arikunto, “Prosedur Penelitian suatu pendekatan praktik,” *Jakarta: Rineka Cipta*, 2010.
- [11] F. Marisa, “Educational Data Mining (Konsep dan Penerapan),” *J. Teknol. Inf. Teor. Konsep, dan Implementasi*, vol. 4, no. 2, 2013, pp. 90–97.