

# Analisis Penderita Covid19 di Indonesia dengan Metode Linier Regresi dan *Unsupervised Learning*

Yana Cahyana<sup>1</sup>, Amril Mutoi Siregar<sup>2</sup>

<sup>1,2</sup>Department of Informatic, Universitas Buana Perjuangan Karawang, Indonesia

---

## Article Info

### Article history:

Received Sep 3, 2021  
Revised Sep 19, 2021  
Accepted Sep 29, 2021

---

### Keywords:

*Covid-19*  
*Regresi*  
*Unsupervised learning*  
*Prediction*  
*k-means*

---

## ABSTRACT

The Covid-19 disease has now been declared a pandemic disease because the level of spread and the risk posed is very dangerous. Various steps such as awareness programs, social distancing, and contact tracing have been taken to control the COVID-19 outbreak. In the absence of a vaccine, prediction of confirmed cases, deaths, and recoveries is needed to increase the capacity of the health care system and control transmission. In this study, cumulative and daily cases were confirmed, died, and recovered in Indonesia. The analysis does not consider any changes in government control measures. Information from this study can provide relevant information to government and health officials and the public. How is the cure rate to the confirmed, the death rate to the number of sufferers? This study uses regression and clustering models with K-means, using unsupervised learning and supervised learning to build the distribution model. The results of this study using the regression method with  $R^2 = 0.99$  while for clustering with  $K = 10 - 15$  intervals seen from the results of the elbow method.

---

## Corresponding Author:

**Amril Mutoi Siregar**

Department of Informatic,  
Universitas Buana Perjuangan Karawang, Indonesia.  
Jl. H.S. Ronggowaluyo, Telukjambe Timur, Karawang 41361.  
Email: amrilmutoi@ubpkarawang.ac.id

---

## 1. PENDAHULUAN

Dunia sedang menghadapi wabah virus korona yang merupakan penyakit menular dan juga dinyatakan sebagai darurat kesehatan masyarakat global oleh Organisasi Kesehatan Dunia (WHO) [1]. Itu berasal dari Wuhan, Hubei, Cina pada Desember 2019 dan menyebar ke seluruh dunia. Organisasi Kesehatan Dunia secara resmi menamai penyakit itu COVID-19 dan mengklasifikasikan wabah COVID-19 sebagai pandemi pada 11 Maret 2020 [2], [3] Virus ini sangat mudah menular.

Salah satu langkah awal yang dilakukan oleh pemerintah yaitu mensosialisasikan gerakan *Social Distancing* untuk masyarakat. Langkah ini bertujuan untuk memutus mata rantai penularan pandemi *covid-19* ini karena langkah tersebut mengharuskan masyarakat menjaga jarak aman dengan manusia lainnya minimal 2 meter, tidak melakukan kontak langsung dengan orang lain serta menghindari pertemuan massal [4]. Namun, pada kenyataannya langkah-langkah tersebut tidak disikapi dengan baik oleh masyarakat, sehingga jumlah kasus terus meningkat. Di samping itu, pelayanan kesehatan di Indonesia dan SDM Kesehatan yang ada dalam menangani kasus *pandemic covid-19* ini juga belum memadai sedangkan kasus terus melonjak naik. Berdasarkan latar belakang dari tulisan ini maka penulis ingin melihat bagaimana Indonesia dalam menghadapi pandemi *covid-19* yang terjadi saat ini. Peneliti sebelumnya telah melakukan beberapa jenis analisa data covid 19, bagaimana merepresentasikan penderita COVID-19 dengan model time series berbasis model pertumbuhan [5], [6].

Dalam skenario yang suram seperti itu, banyak hasil yang berpotensi menarik bagi pembuat kebijakan; misalnya: berapa banyak kasus yang dikonfirmasi, meninggal, dan dipulihkan. Kapan mulai tren penurunan COVID-19. Pemerintah dapat mengambil langkah optimal untuk memperbaiki system kesehatan untuk memperlambat laju penularan dengan mengetahui prediksi dan pengelompokan daerah. Berbagai strategi akan sangat diperlukan untuk menangani wabah saat ini, termasuk pemodelan komputasi, alat statistik, dan

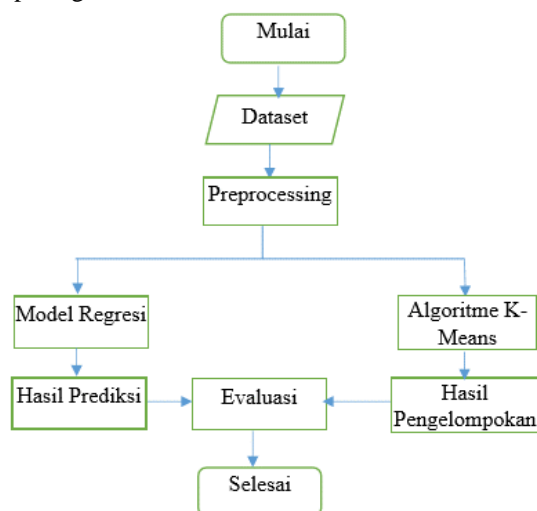
analisis kuantitatif untuk mengontrol penyebaran dan perkembangan pesat dari pengobatan baru [7][8]. Peneliti fokus pada jumlah kasus yang dikonfirmasi, meninggal, dan pulih sebagai target prediksi Peneliti karena terbatasnya data yang tersedia pada hasil lainnya[9].

Beberapa model matematika dan statistik dapat dibuat untuk mempelajari dan menganalisis proses penularan COVID-19. Sehingga peneliti dapat secara akurat memprediksi prevalensi dan mengeksplorasi situasi kemungkinan kasus dan kesembuhan atau kematian. Oleh karena itu, untuk mengurangi bahaya wabah COVID-19, analisis dan penelitian model prediksi Covid-19 menjadi topik penelitian yang hangat [10]. Model yang paling umum digunakan untuk memprediksi penyakit menular adalah prediksi penyakit menular berbasis internet, model prediksi deret waktu, dan model prediksi persamaan diferensial berdasarkan dinamika.

Penelitian ini berfokus pada analisis situasi *pandemic Covid-19* di Indonesia dengan menggunakan model regresi dan metode *unsupervised learning* (algoritma K-means), yang sesuai dengan hukum *statistic epidemiologi*. Pendekatan K-means digunakan untuk pengelompokan provinsi tingkat terkonfirmasi, meninggal dan pulih. Menggunakan algoritma K-means, apakah algoritma tersebut efektif dan model regresi memiliki tujuan menguji data Covid-19 untuk data yang bersifat regresi. Penelitian sebelum hanya menganalisa menggunakan satu macam model analisis seperti regresi, *supervised learning* dan *unsupervised learning*. Penelitian ini menggunakan sekaligus 2 model analisa.

## 2. METODE

Tujuan penelitian ini adalah untuk membuat prediksi dari kasus kumulatif yaitu kasus terkonfirmasi, meninggal, dan pulih setiap hari di Indonesia, menggunakan dua metode yaitu regresi linier dan klastering K-Means. Metode penelitian berisi langkah-langkah yang akan digunakan dalam penelitian ini agar terstruktur dengan baik, sistematisa ini menyebabkan proses penelitian dapat dipahami dan diikuti oleh pihak pembaca. Penelitian yang akan dilakukan adalah untuk menganalisa data yang diperoleh dari pengamatan data yang ada. Langkah yang dilakukan pada penelitian sebelumnya adalah dapat digambarkan pada kerangka berpikir penelitian yang dapat dilihat pada gambar dibawah ini .



Gambar 1. Kerangka berpikir penelitian

### Dataset

Dataset yang digunakan dalam penelitian ini adalah data yang tersedia di Kaggle dan dimana data tersebut kumpulan data penderita COVID 19 yang terjadi di Indonesia, data yang dikumpulkan kejadian harian untuk dijadikan data time series. Dataset yang digunakan mulai tanggal 18 maret 2020 sampai 10 Januari 2021. Dataset untuk regresi linier menggunakan time series dan algoritma k-means dengan mengelompokan provinsi di Indonesia.

### Preprocessing

Data yang dikumpulkan dalam kondisi yang banyak noise sehingga dibutuhkan preprocessing data berupa *cleaning* data, reduksi atribut, ada *preprocessing* terdapat dua tahap, yaitu sebagai berikut: a. Data Cleaning yaitu menghilangkan data yang tidak diperlukan seperti menangani *missing value*, *noise* data serta menangani data yang tidak konsisten dan relevan. b. *Data Integration*, dilakukan terhadap atribut yang mengidentifikasi entitas yang unik

### Model regresi

Tahap ini pengujian model regresi linier terhadap data, untuk melakukan komputasi untuk didapatkan model yang akan dijadikan model prediksi. Dua variabel tersebut adalah variabel dependen (yyy) atau disebut juga dengan variabel respon dan variabel *independent* (xxx) atau disebut juga dengan variabel prediktor dan *variabel* penjelas. Skala data yang digunakan dalam regresi linier sederhana adalah interval atau rasio [3]. Jika terdapat pasangan data  $(x_1, y_1), (x_2, y_2), \dots, \dots, \dots, (x_n, y_n), (x_n, y_n)$ , maka hubungan fungsional pasangan data tersebut dijelaskan dalam model regresi linier sederhana sebagai berikut.

$$\hat{y}_i = b_0 + b_1 x_i \quad 1$$

#### Algoritma K-Means

Pada fase ini dilakukan pemilihan model yang akan digunakan untuk melakukan pengelompokan propinsi diindonesia berdasarkan terinfeksi covid-19. Model atau metode yang akan digunakan pada penelitian ini adalah metode K-Means. Cara kerja algoritma K-Means adalah sebagai berikut. Penerapan K-Means *Clustering* ini dapat dilakukan dengan prosedur step by step berikut [11]:

1. Siapkan data training berbentuk *vector*.
2. Set nilai K cluster.
3. Set nilai awal *centroids*.
4. Hitung jarak antara data dan centroid menggunakan rumus *Euclidean Distance*.
5. Partisi data berdasarkan nilai *minimum*.
6. Kemudian lakukan iterasi selama partisi data masih bergerak (tidak ada lagi objek yang bergerak ke partisi lain), bila masih maka ke poin 3.
7. Bila grup data sekarang sama dengan grup data sebelumnya, maka hentikan iterasi.
8. Data telah dipartisi sesuai nilai centroid akhir.

Berikut ini adalah rumus dari *Euclidean Distance*:

$$(x, y), (a, b) = \sqrt{(x - a)^2 + (y - b)^2} \quad 2$$

Keterangan:

$x$  = atribut x

$y$  = atribut y

$a$  = titik *centroid* a

$b$  = titik *centroid* b.

#### Hasil prediksi

Tahap ini untuk menampilkan hasil prediksi, hasil dari model yang digunakan. Jika hasilnya sesuai dengan harapan dapat selesai pada tahap ini, jika belum sesuai hasil bisa dilakukan dengan melihat dataset yang digunakan.

#### Hasil pengelompokan

Tahap ini untuk menampilkan hasil pengelompokan, hasil dari model yang digunakan. Jika hasilnya sesuai dengan harapan dapat selesai pada tahap ini, jika belum sesuai hasil bisa dilakukan dengan melihat dataset yang digunakan. Hasil akan dikelompokkan ke dalam 3 kelompok yaitu tinggi, sedang dan rendah.

#### Evaluasi

Pengujian akan dilakukan dengan membandingkan pengelompokan yang dilakukan oleh algoritma K-Means dengan pengelompokan yang dilakukan. Mengidentifikasi pola – pola yang menarik dalam *knowledge base* yang diidentifikasi. Pada tahap ini, menghasilkan pola – pola khas maupun model prediksi yang dievaluasi untuk menilai kajian yang ada sudah memenuhi target yang diinginkan. Metode evaluasi yang digunakan SSE dan RMSE dan *elbow metode*.

### 3. HASIL DAN PEMBAHASAN

Setelah melakukan penelitian dengan beberapa tahapan, maka beberapa hasil dapat dilampirkan dalam bentuk grafik dan tabel, agar lebih mudah dipahami. Hasil penelitian metode prediksi dan pengelompokan untuk menganalisa penderita penyakit covid-19 di Indonesia.

#### 3.1. Hasil *preprocessing data*

Data yang dikumpulkan dari Kaggle.com memiliki 17 atribut dan 392 *record* untuk prediksi menggunakan algoritma regresi, setelah melakukan preprocessing data terseleksi menjadi 4 atribut dan 392 *record*, pada tabel 1. Contoh *dataset* yang digunakan untuk prediksi. Analisa menggunakan pengelompokan daerah yang potensial dibuat dalam satu kelompok dengan metode algoritma K-means, dimana memiliki 17 atribut dan 392 *record*, setelah dilakukan seleksi data menjadi 4 atribut dengan 34 *record*, dan pengelompokan ini berdasarkan jumlah propinsi yang ada di Indonesia. Pada tabel 2. Contoh data yang digunakan.

Tabel 1. Contoh dataset hasil preprocessing untuk prediksi

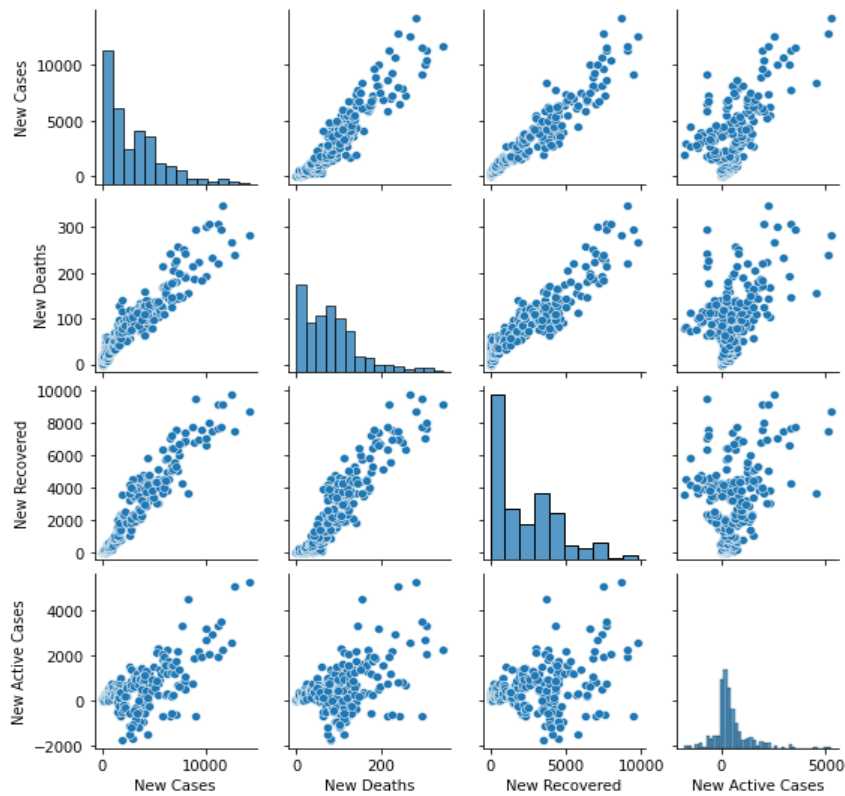
Date	New Cases	New Deaths	New Recovered	New Active Cases
3/2/2020	2	0	0	2
3/3/2020	0	0	0	0
3/4/2020	0	0	0	0
3/5/2020	0	0	0	0
3/6/2020	2	0	0	2
3/7/2020	0	0	0	0
.....	.....	.....	.....	.....
.....	.....	.....	.....	.....
.....	.....	.....	.....	.....
3/24/2021	5227	118	7622	-2513
3/25/2021	6107	98	4656	1353
3/26/2021	4982	85	5679	-782
3/27/2021	4461	198	4243	20
3/28/2021	4083	85	4279	-281

Tabel 1 dan tabel 2 merupakan jenis data yang digunakan. Terdiri dari hasil seleksi data yang diperoleh. Penelitian ini menganalisis dengan 2 metode yaitu metode prediksi menggunakan algoritma regresi dan pengelompokan menggunakan algoritma K-means.

Tabel 2. Contoh dataset hasil preprocessing untuk pengelompokan

Province	Daily_Case	Daily_Death	Daily_Recov- ered	Ac- tive_Case
Aceh	11237	449	9785	1003
Bali	44896	1231	41985	1680
Banten	47451	1207	44197	2047
Bengkulu	6959	173	5939	847
Diy	39824	965	35045	3814
Dki	410363	6671	396605	7087
Gorontalo	5400	165	5080	155
Jambi	7808	115	6248	1445
Jabar	282631	3753	248235	30643
Jateng	184620	8236	166003	10381
Jatim	148183	10706	135279	2198
Kalimantan barat	7931	43	6900	988
Kalimantan selatan	33058	949	30001	2108
Kalimantan tengah	20228	453	16996	2779
Kalimantan timur	68817	1640	65519	1658
Kalimantan utara	11786	186	11035	565

Kepulauan bangka belitung	13739	217	11836	1686
Kepulauan riau	11500	257	9769	1474
Lampung	16116	816	13948	1352
Maluku	7535	118	7137	280
Maluku utara	4417	120	4147	150
Nusa tenggara barat	10488	389	7879	2220
Nusa tenggara timur	14636	376	12736	1524
Papua	20365	207	11315	8843
Papua barat	8986	150	8550	286
Riau	45237	1115	39006	5116
Sulawesi barat	5472	119	5259	94
Sulawesi selatan	61524	934	60288	302
Sulawesi tengah	12316	335	11225	756
Sulawesi tenggara	10427	210	9886	331
Sulawesi utara	15657	507	13240	1910
Sumatera barat	37380	809	34144	2427
Sumatera selatan	20725	1006	18131	1588
Sumatera utara	29522	976	26250	2296

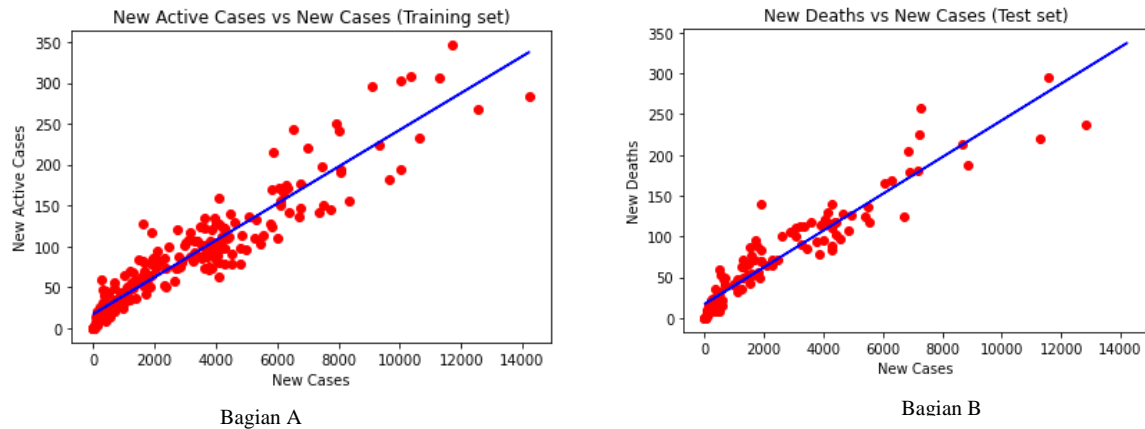


Gambar 2. Ploting data yang digunakan

Gambar 2. berfungsi melihat secara visual data yang akan diteliti, agar pendekatan yang digunakan tepat dengan data yang dimiliki. Tampilan secara visual di Gambar atas memiliki 4 atribut yang digunakan dalam penelitian ini.

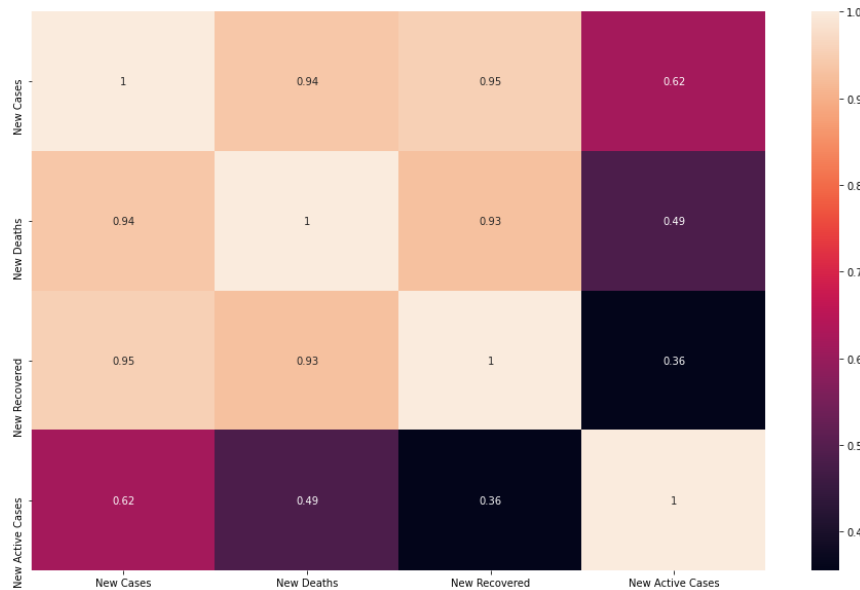
### 3.1 Hasil prediksi

Hasil penelitian menggunakan metode regresi dalam bentuk visual seperti grafik yang memiliki garis hyperplane secara linier. Berikut dibawah ini hasil model regresinya.



Gambar 3. Hasil prediksi dengan metode regresi

Gambar 3 bagian A hasil model *training set* merupakan ilustrasi hasil prediksi antara kasus baru (*new cases*) dengan kasus aktif baru (*new active cases*). Hasilnya dapat diambil kesimpulan bahwa diantara 2 variabel atau atribut tersebut sangat berkorelasi kuat ketika kasus baru dari *interval* 0-6000 sedangkan semakin tinggi kasus baru semakin lemah korelasi diantara dua atribut tersebut. Bagian B hasil model *test set* merupakan ilustrasi hasil prediksi antara kasus baru (*new cases*) dengan kasus meninggal baru (*new deaths*). Hasilnya dapat diambil kesimpulan bahwa diantara 2 variabel atau atribut tersebut sangat berkorelasi kuat ketika kasus baru (*new cases*) dari *interval* 0-6000 sedangkan semakin tinggi kasus baru (*new cases*) semakin lemah korelasi diantara dua atribut tersebut. Jadi dapat diambil kesimpulan atribut kasus baru (*new cases*), kasus aktif baru (*new active cases*), dan meninggal baru (*new deaths*) sangat berkorelasi.



Gambar 4. korelasi data terhadap atribut

Gambar 4 menampilkan korelasi data diantara atribut, dapat diambil simpulan bahwa ada beberapa data korelasinya sangat tinggi dan ada juga yang lemah. Korelasi tinggi adalah atribut sembuh baru (*new recovered*) dengan kasus baru (*new cases*) sebesar 0.95. dan kasus baru (*new cases*) dengan meninggal baru (*new deaths*) serta meninggal baru (*new deaths*) dengan kasus baru (*new cases*) sebesar 0.94. korelasi lemah adalah atribut kasus aktif baru (*new active cases*) dengan sembuh baru (*new recovered*) dan sembuh baru (*new recovered*) dengan kasus aktif baru (*new active cases*) sebesar 0.36.

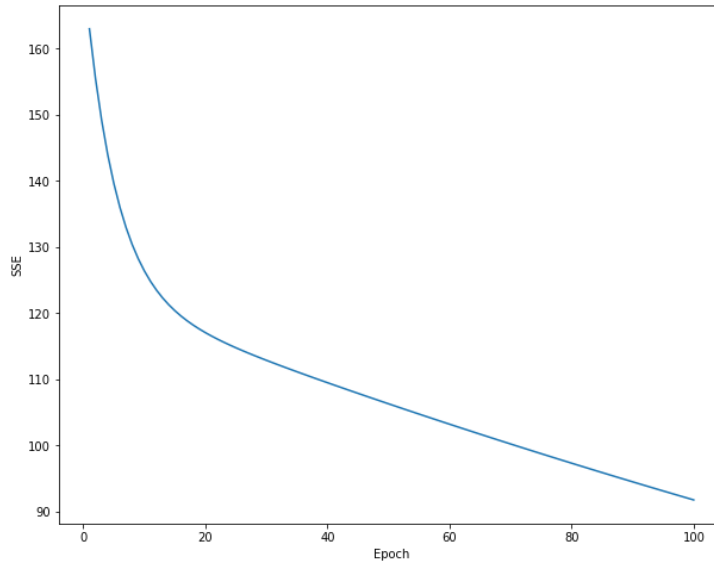
Hasil penelitian ini diukur menggunakan RMSE dan R2 dengan hasil sebagai berikut.

*The model performance of training set*

-----  
 RMSE of training set is 2.8176745352933008e-09  
 R2 score of training set is 1.0

*The model performance of testing set*

-----  
 RMSE of testing set is 6.930601000425868e-05  
 R2 score of testing set is 0.9999999999999911

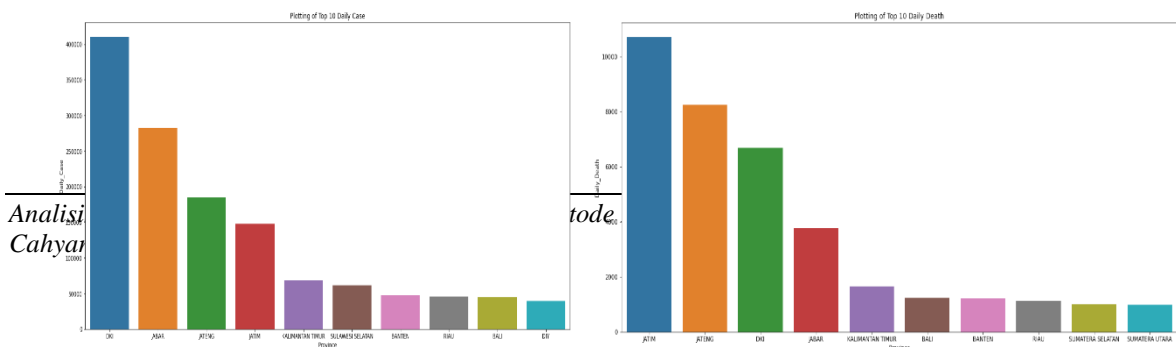


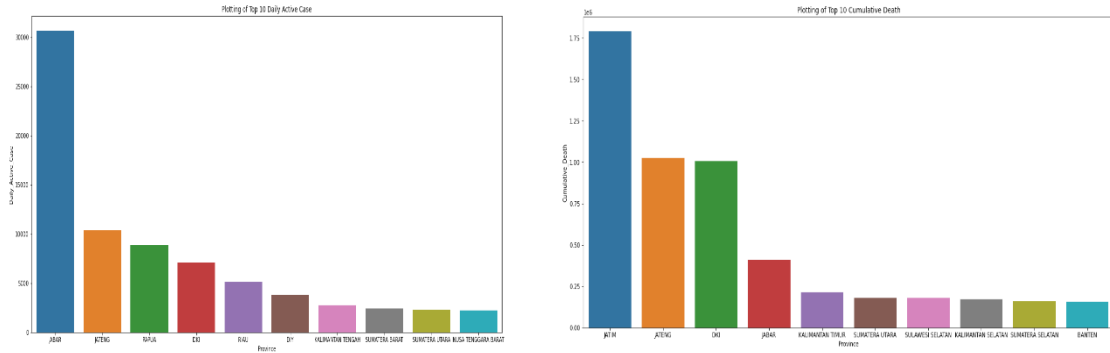
Gambar 5. proyeksi hasil prediksi

Gambar 5 merupakan hasil prediksi menggunakan grafik dimana dapat disimpulkan bahwa semakin kecil *epoch* maka semakin tinggi *Sum square error (SSE)* begitu juga semakin besar epochnya semakin kecil *Sum square error (SSE)*.

3.2 Hasil *unsupervised learning*

. Gambar 6 merupakan hasil dari TOP 10 propinsi terhadap penderita kasus Covid 19 di Indonesia. Grafik tersebut dapat diambil simpulan bahwa kasus harian masih didominasi provinsi DKI, Jabar, Jateng dan Jatim. Kasus meninggal baru didominasi Jatim, Jateng, DKI dan Jabar. Kasus aktif baru didominasi Jabar, Jateng, Papua, dan DKI. Kumulatif meninggal di dominasi Jatim, Jateng, DKI, dan Jabar.

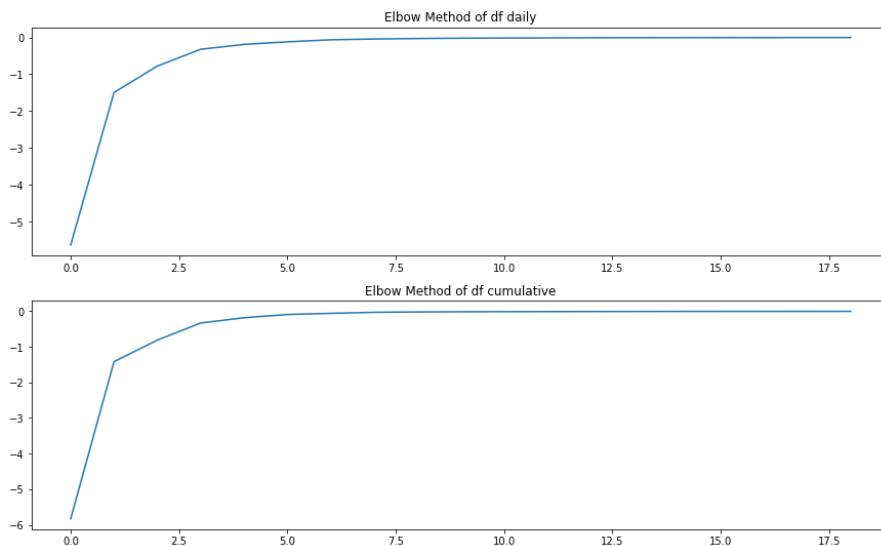




Gambar 6. Hasil *plotting* top 10 pada provinsi

Simpulan dalam dataset menjelaskan kepada pembaca bahwa data ini adalah tentang kasus harian, kematian, sembuh, kelas aktif di Indonesia. Data ini menjelaskan kepada kita tentang masing-masing provinsi seberapa tinggi kasus, kematian, sembuh, kelas aktif. Pada dataset ini diketahui kasus harian tertinggi adalah DKI, kematian harian Jatim, harian sembuh DKI, kasus aktif harian Jabar. Pada dataset ini terdapat fitur kumulatif, dan kumulatif kasus tertinggi adalah DKI, kumulatif kematian Jatim, kumulatif sembuh DKI, kumulatif kasus aktif Jateng.

Hasil penelitian menggunakan metode *unsupervised learning* dalam bentuk visual seperti grafik yang memiliki garis. Berikut dibawah ini hasil algoritma K-means.



Gambar 7. Hasil pengukuran Metode *Elbow*

Gambar 7 menunjukkan bahwa K yang paling baik dalam penelitian ini adalah interval 10-15 dan semakin banyak nilai K yang digunakan tidak mempengaruhi hasilnya. Gambar 7 merupakan hasil *clustering* dengan kumulatif kasus dan harian. Pada tabel 3 merupakan hasil pengelompokan dengan k=2. Keterangannya 0 = kelompok rendah dan 1 = kelompok tinggi. Kumulatif dengan *daily* (harian) sama *trend* hasil visualisasi menggunakan *elbow*.

Tabel 3. Hasil Pengelompokan berdasarkan provinsi

Province	Cluster
Aceh	0
Bali	0
Banten	0
Bengkulu	0



Diy	0
Gorontalo	0
Jambi	0
Jatim	1
Kalimantan barat	0
Kalimantan selatan	0
Kalimantan tengah	0
Kalimantan timur	0
Kalimantan utara	0
Kepulauan bangka belitung	0
Kepulauan riau	0
Lampung	0
Maluku	0
Maluku utara	0
Nusa tenggara barat	0
Nusa tenggara timur	0
Papua	0
Papua barat	0
Riau	0
Sulawesi barat	0
Sulawesi selatan	0
Sulawesi tengah	0
Sulawesi tenggara	0
Sulawesi utara	0
Sumatera barat	0
Sumatera selatan	0
Sumatera utara	0
Dki	1
Jabar	1
Jateng	1

Berdasarkan tabel 3. pengelompokan diketahui ada empat provinsi dalam fitur harian yang termasuk ke dalam *cluster* 1 adalah ['DKI', 'Jabar', 'Jateng', 'Jatim'] dan sisanya adalah provinsi yang termasuk dalam *cluster* 0 seperti ['Aceh', 'Bali', 'Banten', 'Bengkulu'], serta empat provinsi dalam fitur kumulatif yang termasuk dalam *cluster* 1 adalah ['Dki', 'Jabar', 'Jateng', 'Jatim'] dan provinsi lainnya termasuk dalam cluster 0 seperti ['Aceh', 'Bali', 'Banten', 'Bengkulu']. Penelitian ini dapat diambil kesimpulan bahwa tidak semua provinsi ketika lebih tinggi kasus baru akan semakin tinggi yang meninggal baru, berarti provinsi penanganan pasien *Covid-19* yang optimal akan mengurangi tingkat kematian.

#### 4. PENUTUP

Berdasarkan hasil penelitian ini dapat diambil beberapa kesimpulan yaitu.

1. Menggunakan metode regresi untuk prediksi dalam kasus Analisa penderita *Covid-19* sangat baik dimana hasilnya  $r^2 = 0,99$ , jadi Analisa penderita sangat baik digunakan untuk model.
2. Metode *clustering* sebagai *unsupervised learning* dapat digunakan untuk Analisa data dengan pengelompokan provinsi berdasarkan kesamaan data dimana K yang paling maksimal di *interval* 10-15, sehingga model Analisa pengelompokan sangat baik dan mudah dipahami.

3. Penelitian berikut diharapkan dapat membuat API yang berbasis Android agar penelitian ini berguna bagi masyarakat.

#### UCAPAN TERIMAKASIH

Terimakasih kasih kepada Direktorat Jenderal Pendidikan Tinggi Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi atas Hibah Tahun anggaran 2021 dan kampus UBK Karawang yang memberikan semangat sehingga penelitian ini dapat berjalan tepat waktu.

#### DAFTAR PUSTAKA

- [1] L. Wang, Y. Wang, D. Ye, and Q. Liu, "Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence," *Int. J. Antimicrob. Agents*, vol. 55, no. 6, p. 105948, Jun. 2020, doi: 10.1016/j.ijantimicag.2020.105948.
- [2] F. R. Pratikto, "Prediksi Akhir Pandemi COVID-19 di Indonesia dengan Simulasi Berbasis Model Pertumbuhan Parametrik," *J. Rekayasa Sist. Ind.*, vol. 9, no. 2, pp. 63–68, Jul. 2020, doi: 10.26593/jrsi.v9i2.4018.63-68.
- [3] K. Sarkar, S. Khajanchi, and J. J. Nieto, "Modeling and forecasting the COVID-19 pandemic in India," *Chaos, Solitons & Fractals*, vol. 139, p. 110049, Oct. 2020, doi: 10.1016/j.chaos.2020.110049.
- [4] D. R. Buana, "Analisis Perilaku Masyarakat Indonesia dalam Menghadapi Pandemi Virus Corona (Covid-19) dan Kiat Menjaga Kesejahteraan Jiwa," *SALAM J. Sos. dan Budaya Syar-i*, vol. 7, no. 3, Mar. 2020, doi: 10.15408/sjsbs.v7i3.15082.
- [5] N. Nuraini, K. Khairudin, and M. Apri, "Modeling Simulation of COVID-19 in Indonesia based on Early Endemic Data," *Commun. Biomath. Sci.*, vol. 3, no. 1, pp. 1–8, Apr. 2020, doi: 10.5614/cbms.2020.3.1.1.
- [6] C. Y. Shen, "Logistic growth modelling of COVID-19 proliferation in China and its international implications," *Int. J. Infect. Dis.*, vol. 96, pp. 582–589, Jul. 2020, doi: 10.1016/j.ijid.2020.04.085.
- [7] Y. Dong *et al.*, "Epidemiology of COVID-19 Among Children in China," *Pediatrics*, vol. 145, no. 6, p. e20200702, Jun. 2020, doi: 10.1542/peds.2020-0702.
- [8] R. Tosepu *et al.*, "Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia," *Sci. Total Environ.*, vol. 725, p. 138436, Jul. 2020, doi: 10.1016/j.scitotenv.2020.138436.
- [9] Y. Fang, Y. Nie, and M. Penny, "Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis," *J. Med. Virol.*, vol. 92, no. 6, pp. 645–659, Jun. 2020, doi: 10.1002/jmv.25750.
- [10] P. Sarkar, N. Debnath, and D. Reang, "Coupled human-environment system amid COVID-19 crisis: A conceptual model to understand the nexus," *Sci. Total Environ.*, vol. 753, p. 141757, Jan. 2021, doi: 10.1016/j.scitotenv.2020.141757.
- [11] N. Nidheesh, K. A. Abdul Nazeer, and P. M. Ameer, "An enhanced deterministic K-Means clustering algorithm for cancer subtype prediction from gene expression data," *Comput. Biol. Med.*, vol. 91, pp. 213–221, Dec. 2017, doi: 10.1016/j.compbimed.2017.10.014.