

# Estimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Simalungun

Fica Oktavia Lusiana\*, Indri Fatma, Agus Perdana Windarto

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: <sup>1</sup>\*caalucyana@gmail.com, <sup>2</sup>indriifatmaa99@email.com

**Abstrak**—Lembaga Badan Pusat Statistik (BPS) merupakan lembaga pemerintah non-Departemen yang bertanggung jawab langsung kepada presiden. Penggunaan dari pengumpulan data tersebut adalah guna pendataan negara untuk kebutuhan strategi ekonomi, infrastruktur, dan sebagainya. Sehingga lembaga BPS harus dapat memprediksi laju estimasi pertumbuhan penduduk. Terutama salah satu lembaga BPS di Sumatera Utara ialah BPS Simalungun mengalami kendala dalam mengestimasi laju pertumbuhan penduduk. Model regresi linier berganda merupakan pengembangan dari model regresi linier sederhana. Jika pada model regresi linier sederhana hanya terdiri dari satu variabel bebas dan satu variabel terikat, maka pada model regresi linier berganda jumlah variabel bebasnya lebih dari satu dan satu variabel terikat. Tahapan yang dilakukan pada proses data mining diawali dari seleksi data dari data sumber ke data target, tahap preprocessing untuk memperbaiki kualitas data, transformasi, data mining serta tahap interpretasi dan evaluasi yang menghasilkan output berupa pengetahuan baru yang diharapkan memberikan kontribusi yang lebih baik.

**Kata Kunci:** Badan Pusat Statistik; Regresi Linear Berganda; Data Mining; Preprocessing; Transformasi

**Abstract**—The Central Statistics Agency (CSA) is a non-departmental government agency that reports directly to the president. The use of data collection is for state data collection for the needs of economic strategies, infrastructure, and so on. So that the CSA institution must be able to predict the estimated rate of population growth. Particularly one of the CSA institutions in North Sumatra, CSA Simalungun, has experienced problems in estimating the population growth rate. Multiple linear regression model is the development of a simple linear regression model. If the simple linear regression model consists of only one independent variable and one dependent variable, then in multiple linear regression the number of independent variables is more than one and one dependent variable. The stages carried out in the data mining process begin with data selection from source data to target data, the preprocessing stage to improve data quality, transformation, data mining and interpretation and evaluation stages which produce output in the form of new knowledge which is expected to make a better contribution.

**Keywords:** Central Statistics Agency; Multiple Linear Regression; Data Mining; Preprocessing; Transformation

## 1. PENDAHULUAN

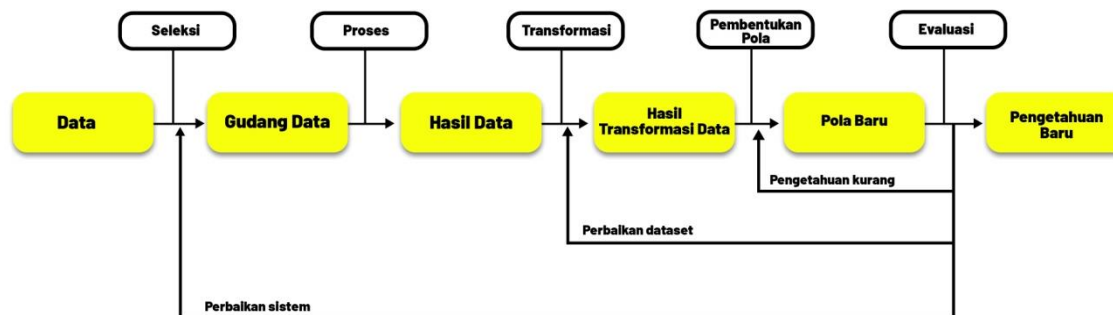
Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses Knowledge Discovery in Database (KDD) secara keseluruhan [1]. Berbagai macam metode-metode yang terdapat pada data mining [2]–[5], contohnya adalah metode estimasi [6]. Metode estimasi bekerja ketika himpunan data sebagai sampel data bersifat numerik dan memiliki label. Metode ini juga digunakan untuk menerka sebuah nilai yang belum diketahui, contoh kasus yang saat ini diambil adalah laju pertumbuhan penduduk.

Badan Pusat Statistik (BPS) mempunyai tugas dan fungsi melakukan pendataan statistik terhadap jumlah penduduk dari tahun ke tahun. Lembaga BPS merupakan lembaga pemerintah non-Departemen yang bertanggung jawab langsung kepada presiden. Penggunaan dari pengumpulan data tersebut adalah guna pendataan negara untuk kebutuhan strategi ekonomi, infrastruktur, dan sebagainya. Sehingga lembaga BPS harus dapat memprediksi laju estimasi pertumbuhan penduduk. Maka dari itu penggunaan proses data mining pada data penduduk BPS dapat mengestimasi laju pertumbuhan penduduk. Terutama salah satu lembaga BPS di Sumatera Utara ialah BPS Simalungun mengalami kendala dalam mengestimasi laju pertumbuhan penduduk. Sehingga solusi ini dari permasalahan ini dapat diselesaikan menggunakan metode data mining dengan metode diantaranya metode estimasi menggunakan teknik Regresi Linier Berganda. Selanjutnya yang diharapkan dari teknik regresi linier berganda dapat membangun sebuah perangkat lunak yang memudahkan dalam mengestimasi laju pertumbuhan penduduk.

## 2. METODOLOGI PENELITIAN

### 2.1 Data Mining

Data mining merupakan suatu langkah dalam *Knowledge Discovery in Databases* (KDD). Knowledge discovery sebagai suatu proses terdiri atas pembersihan data (*data cleaning*), integrasi data (*data integration*), pemilihan data (*data selection*), transformasi data (*data transformation*), data mining, evaluasi pola (*pattern evaluation*) dan penyajian pengetahuan (*knowledge presentation*). Data mining mengacu pada proses untuk menambang (mining) pengetahuan dari sekumpulan data yang sangat besar [7]. Tahapan yang dilakukan pada proses data mining diawali dari seleksi data dari data sumber ke data target, tahap preprocessing untuk memperbaiki kualitas data, transformasi, data mining serta tahap interpretasi dan evaluasi yang menghasilkan output berupa pengetahuan baru yang diharapkan memberikan kontribusi yang lebih baik.

Gambar 1. Tahapan *Data Mining*

## 2.2. Metode Regresi Linear

Dalam menggunakan analisis regresi linier, syarat pertama yang harus adalah data yang akan digunakan harus berskala interval dan Rasio. Akan tetapi jika data yang berkala ordinal tetapi dipaksakan menggunakan analisis regresi linier, maka akan diperoleh koefisien korelasi yang kecil dan tidak dapat memenuhi syarat sebagaimana yang diharuskan dalam model regresi yaitu model fit sehingga peneliti akan keliru dalam melakukan interpretasi dari model regresi tersebut. Untuk mengatasi masalah tersebut maka terdapat 2 solusi yang dapat dilakukan, pertama dengan tetap menggunakan data ordinal dan mengganti teknik analisis datanya, atau mengganti/mentransformasi data dari nominal menjadi interval agar analisis regresi linier dapat digunakan[8].

Model regresi linier berganda merupakan pengembangan dari model regresi linier sederhana. Jika pada model regresi linier sederhana hanya terdiri dari satu variabel bebas dan satu varuiabel terikat, maka pada regresi linier berganda jumlah variabel bebasnya lebih dari satu dan satu variabel terikat. Dengan bertambahnya variabel variabel bebas maka bentuk umum dari persamaan regresi linier berganda yang mencakup dua atau lebih variabel bebas adalah sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (1)$$

Apabila menggunakan data sampel, maka dapat menggunakan model estimasi dari Persamaan (1) adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + e \quad (2)$$

Berdasarkan Persamaan (2) maka bentuk matriks dari model regresi linier berganda dapat dituliskan sebagai berikut:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{k1} \\ 1 & X_{12} & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad (3)$$

atau dapat ditulis dengan  $Y = X \beta + e$ , dengan:

$Y$  adalah vektor pengamatan berukuran  $n \times 1$

$X$  adalah matriks variabel bebas ukuran  $n \times k$

$\beta$  adalah vektor parameter yang akan ditaksir berukuran  $k \times 1$

$e$  adalah vektor random error berukuran  $n \times 1$

## 3. HASIL DAN PEMBAHASAN

Data mining berasal dari permasalahan ledakan data (*data explosion*) yang dialami oleh instansi/perusahaan yang telah mengumpulkan data dari berbagai macam transaksi. Contoh nya adalah data pembelian, data nasabah, data transaksi, data penjualan, dan masih banyak lagi. Data tersebut akan “menggungung” jika dibiarkan begitu saja, layaknya sebuah barang milik kita yang kita biarkan menumpuk semakin besar, apakah barang tersebut kita buang begitu saja atau kita bisa “menambang”-nya untuk dicari barang atau informasi penting yang masih sangat berguna. Berikut juga data-data para penduduk yang tersimpan di *database* BPS Simalungun yang berdiri sudah puluhan tahun. Memiliki segudang data yang sangat banyak sekali, maka dari itu permasalahan ini bisa dimanfaatkan untuk mengetahui estimasi laju pertumbuhan pendudukan dengan data-data yang ada pada BPS Simalungun. Dengan penerapan metode linear regresi berganda yang sudah populer untuk mengetahui estimasi atau prediksi maka penerapan metode ini sangat cocok untuk digunakan. Berikut beberapa tahapan penyelesaian masalah dengan metode regresi linear berganda :

a) Mempersiapkan *data training*, yaitu data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan dalam kelas-kelas tertentu.

b) Menentukan variabel bebas dan variabel tidak bebas

Variabel bebas : - Jumlah Laki-Laki (X1)

- Jumlah Perempuan (X2)

Variabel tidak bebas : Jumlah Penduduk (Y)

c) Mencari nilai persamaan

$$Y = a + b_1.x_1 + b_2.x_2$$

d) Menentukan nilai konstanta dan koefisien regresi

Data yang akan diambil pada BPS Simalungun ialah data selama 2016-2020. Data akan digunakan dalam perhitungan metode regresi linear berganda untuk menganalisa mengenai estimasi pertumbuhan penduduk. Data ini nantinya akan bentuk menjadi sebuah informasi baru untuk tahun mendatang.

**Tabel 1.** Data Penduduk Kabupaten Simalungun

No.	Kecamatan	2016	2017	2018	2019	2020
1.	Silimakuta	15.542	15.777	16.083	16.376	16.656
2.	Pamatang Silimahuta	10.765	10.834	10.898	10.959	11.016
3.	Purba	23.708	24.027	24.325	24.608	24.878
4.	Haranggaol Horison	5.070	5.080	5.090	5.099	5.108
5.	Dolok Pardamean	16.180	16.201	12.931	12.948	12.994
6.	Sidamanik	27.597	27.676	27.750	27.819	27.883
7.	Pamatang Sidamanik	16.611	16.659	16.703	16.745	16.784
8.	Girsang Sipangan Bolon	14.810	14.886	14.956	15.023	15.085
9.	Tanah Jawa	47.508	47.646	47.773	47.892	48.004
10.	Hatonduhan	21.342	21.366	21.389	21.409	21.428
11.	Dolok Panribuan	18.309	18.363	18.411	18.457	18.500
12.	Joriang Hataran	15.623	15.667	15.709	15.747	15.784
13.	Panei	22.095	22.199	22.296	22.386	22.472
14.	Panombeian Panei	19.503	19.547	19.587	19.625	19.660
15.	Raya	32.070	32.260	25.965	26.098	26.191
16.	Dolok Masagal	-	-	9.762	9.796	9.831
17.	Dolok Silou	14.314	14.411	14.500	14.584	14.664
18.	Silou Kahean	17.482	17.555	17.624	17.689	17.749
19.	Raya Kahean	17.775	17.830	17.882	17.931	17.976
20.	Tapian Dolok	40.708	41.154	41.572	41.967	42.342
21.	Dolok Batu Nanggar	40.490	40.663	40.824	40.974	41.116
22.	Siantar	65.833	66.304	66.743	67.156	67.546
23.	Gunung Malela	34.473	34.767	35.042	35.301	35.546
24.	Gunung Maligas	27.675	27.923	28.153	28.370	28.576
25.	Hutabayu Raja	29.722	29.808	29.887	29.961	30.031
26.	Jawa Maraja Bah Jambi	21.719	22.021	22.302	22.570	22.825
27.	Pamatang Bandar	31.640	31.679	31.714	31.747	31.778
28.	Bandar Huluan	26.377	26.473	26.563	26.646	26.725
29.	Bandar	68.187	68.958	69.680	70.360	71.009
30.	Bandar Masilam	24.804	24.876	24.941	25.003	25.061
31.	Bosar Maligas	40.371	40.591	40.797	40.989	41.171
32.	Ujung Padang	41.193	41.288	41.376	41.458	41.534
Total :		849.405	854.489	859.228	863.693	867.922

**Tabel 2.** Akumulasi Data Pendudukan Per Tahun Berdasarkan Jenis Kelamin

Tahun	Penduduk		Jumlah Penduduk Y
	Laki-Laki X1	Perempuan X2	
2016	423.202	426.203	849.405
2017	425.794	428.695	854.489
2018	428.105	431.123	859.228
2019	430.306	433.387	863.693
2020	432.425	435.497	867.922
Jumlah	2.139.832	2.154.905	4.294.737

Berdasarkan data yang sudah didapat, perhitungan regresi linear berganda dilakukan dengan banyak perhitungan perkalian dan perpangkatan maka dari itu penyederhanaan data bilangan bisa dilakukan untuk memudahkan perhitungan.

**Tabel 3.** Penyederhanaan Data Penduduk Dibagi 1000

Tahun	Penduduk		Jumlah Penduduk Y
	Laki-Laki X1	Perempuan X2	
2016	423,202	426,203	849,405
2017	425,794	428,695	854,489
2018	428,105	431,123	859,228
2019	430,306	433,387	863,693
2020	432,425	435,497	867,922
Jumlah	2.139,832	2.154,905	4.294,737

Regresi linear berganda merupakan bentuk hubungan di mana variabel bebas X maupun variabel tergantung Y sebagai faktor berpangkat dua. Persamaan umumnya adalah :

$$Y = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n$$

Keterangan :

Y = variabel terikat/ variabel dependent (nilai yang diprediksi)

$a_0, a_1, a_2, a_n$  = koefisien regresi

$X_1, X_2, \dots, X_n$  = variabel bebas/variabel independent

Estimasi pertumbuhan penduduk sebagai (Y) di BPS Simalungun ditinjau dari 2 variabel yaitu jumlah Laki-Laki (X1) dan Jumlah Perempuan (X2) yang akan diprediksi menggunakan teknik metode regresi linear berganda. Maka langkah selanjutnya adalah mencari nilai konstanta dan variabel regresi setiap variabel bebas.

**Tabel 4.** Learning Dataset

Tahun	Penduduk		Jumlah Penduduk Y
	Laki-Laki X1	Perempuan X2	
2016	423,202	426,203	849,405
2017	425,794	428,695	854,489
2018	428,105	431,123	859,228
2019	430,306	433,387	863,693
2020	432,425	435,497	867,922
Jumlah ( $\Sigma$ )	2.139,832	2.154,905	4.294,737

Proses selanjutnya adalah melakukan ikhtisar perhitungan dari variabel nilai yang sudah terdapat pada tabel diatas.

**Tabel 5.** Ikhtisar Perhitungan

Tahun	$X1^2$	$X1*X2$	$X1*Y$	$X2^2$	$X2*Y$
2016	180.000	180.370	359.470	181.649	362.019
2017	181.300	182.535	363.836	183.779	366.315
2018	183.274	184.566	367.840	185.867	370.433
2019	185.163	186.489	371.652	188.344	374.831
2020	186.991	188.320	375.311	189.657	377.977
Jumlah	916.728	922.280	1.838.109	929.296	1.851.575

Setelah jumlah ikhtisar perhitungan sudah didapatkan, untuk mempermudah perhitungan maka dilakukan pembagian lagi.

**Tabel 6.** Penyederhanaan Ikhtisar Perhitungan Dibagi 10000

Tahun	$X1^2$	$X1*X2$	$X1*Y$	$X2^2$	$X2*Y$
2016	18,000	18,037	35,947	18,164	36,201
2017	18,130	18,253	36,383	18,377	36,631
2018	18,327	18,456	36,784	18,586	37,043
2019	18,516	18,648	37,165	18,834	37,483
2020	18,699	18,832	37,531	18,965	37,797
Jumlah	91,672	92,228	183,810	92,929	185,157

Untuk mendapatkan koefisien regresi  $a, b_1$  dan  $b_2$  dapat dilakukan dengan cara simultan dari tiga persamaan, yaitu :

$$a_n + b_1 \Sigma X_1 + b_2 \Sigma X_2 = \Sigma Y \quad \text{.....} \quad \text{(Persamaan 1)}$$

$$a \Sigma X_1 + b_1 \Sigma X_1^2 + b_2 \Sigma X_1 X_2 = \Sigma X_1 Y \quad \text{.....} \quad \text{(Persamaan 2)}$$

$$a \Sigma X_2 + b_1 \Sigma X_1 X_2 + b_2 \Sigma X_2^2 = \Sigma X_2 Y \quad \text{.....} \quad \text{(Persamaan 3)}$$

Kemudian masukkan angka ikhtisar perhitungan (Tabel 6) dan nilai  $\Sigma X_1$ ,  $\Sigma X_2$  dan  $\Sigma Y$  (Tabel 4)

$$a5 + b_1(2139,832) + b_2(2154,905) = 4294,737 \quad \text{.....} \quad \text{(Persamaan 1)}$$

$$a(2139,832) + b_1(91,6728) + b_2(92,2280) = 183,8109 \quad \text{.....} \quad \text{(Persamaan 2)}$$

$$a(2154,905) + b_1(92,2280) + b_2(92,9296) = 185,1575 \quad \text{.....} \quad \text{(Persamaan 3)}$$

Selanjutnya langkah penyelesaiannya adalah sebagai berikut :

1. Persamaan 1 dan 2 di eliminasi

$$\begin{array}{rcl} 5a + b_1(2139,832) + b_2(2154,905) & = & 4294,737 \quad \quad \quad \times 2139,832 \\ (2139,832)a + b_1(91,6728) + b_2(92,2280) & = & 183,8109 \quad \quad \quad \times 5 \\ 10699,16 a + 4578880,988224 b_1 + 4611134,67596 b_2 & = & 9190015,664184 \\ 10699,16 a + 458,364 b_1 + 461,14 b_2 & = & 919,0545 \\ 4578422,624224 b_1 + 4610673,53596 b_2 & = & 9189096,609684 \quad \text{(Persamaan 4)} \end{array}$$

2. Persamaan 1 dan 3 di eliminasi

$$\begin{array}{rcl} 5a + b_1(2139,832) + b_2(2154,905) & = & 4294,737 \quad \quad \quad \times 2154,905 \\ (2154,905)a + b_1(92,2280) + b_2(92,9296) & = & 185,1575 \quad \quad \quad \times 5 \\ 10774,525 a + 4611134,67596 b_1 + 4643615,559025 b_2 & = & 9254750,234985 \\ 10774,525 a + 461,14 b_1 + 464,648 b_2 & = & 925,7875 \\ 4610673,53596 b_1 + 4643150,911025 b_2 & = & 9253824,447485 \quad \text{(Persamaan 5)} \end{array}$$

3. Persamaan 4 dan 5 di eliminasi

$$\begin{array}{rcl} 4578422,624224 b_1 + 4610673,53596 b_2 & = & 9189096,609684 \quad \times 4578422,624224 \\ 4610673,53596 b_1 + 4643150,911025 b_2 & = & 9253824,447485 \quad \times 4546397,302371927 \\ 20961953726006,18 b_1 + 21109612029950,13 b_2 & = & 42071567813957,28 \\ 20961953726006,18 b_1 + 21109608776389,82 b_2 & = & 42071562504669,19 \\ & & 3253560,31 b_2 = 5309288,09 \\ & & b_2 = 5309288,09 : 3253560,31 \\ & & b_2 = 1,6318 \end{array}$$

4. Lalu untuk mencari nilai  $b_1$  masukkan nilai  $b_2$  pada persamaan 4 dengan substitusi

$$\begin{array}{rcl} 4578422,624224 b_1 + 4610673,53596 b_2 & = & 9189096,609684 \\ 4578422,624224 b_1 + 4610673,53596 (1,6318) & = & 9189096,609684 \\ 4578422,624224 b_1 + 7523697,075979528 & = & 9189096,609684 \\ 4578422,624224 b_1 = 7523697,075979528 + 9189096,609684 & & \\ 4578422,624224 b_1 = 16712793,68566353 & & \\ & & b_1 = 16712793,68566353 : 4578422,624224 \\ & & b_1 = 3,6503 \end{array}$$

5. Masukkan nilai  $b_1$  dan  $b_2$  kedalam persamaan 1 dengan substitusi

$$\begin{array}{l} 5a + b_1(2139,832) + b_2(2154,905) = 4294,737 \\ 5a + (3,6503)(2139,832) + (1,6318)(2154,905) = 4294,737 \\ 5a + 7811,0287496 + 3516,373979 = 4294,737 \\ 5a + 11327,4027286 = 4294,737 \\ 5a = 4294,737 - 11327,4027286 \\ 5a = -7032,6657286 \\ a = -7032,6657286 : 5 \\ a = -1406,533 \end{array}$$

Maka telah didapat nilai  $a$ ,  $b_1$  dan  $b_2$  yaitu :

$$a = -1406,533$$

$$b_1 = 3,6503$$

$$b_2 = 1,6318$$

Sehingga menghasilkan persamaan regresi :

$$Y = -1406,533 + 3,6503 b_1 + 1,6318 b_2$$

Setelah persamaan regresi telah didapatkan, maka langkah selanjutnya untuk mengestimasi laju pertumbuhan penduduk pada tahun 2021 adalah dengan memasukkan nilai  $X_1$  dan  $X_2$  pada periode tahun 2020 yaitu dengan  $X_1$  sebanyak 432,425 dan  $X_2$  435,497. Maka kita gunakan kembali persamaan regresi linear berganda untuk mengestimasi laju pertumbuhan penduduk :

$$Y = a + b_1 X_1 + b_2 X_2$$

$$Y = -1406,533 + 3,6503 (432,425) + 1,6318 (435,497).$$

$$Y = -1406,533 + 1578,4809775 + 710,6440046$$

$$Y = 882,5919 \text{ dikali } 1.000$$

$$Y = 882.591 \text{ Penduduk}$$

Dari hasil tersebut maka estimasi pertumbuhan penduduk di Kabupaten Simalungun untuk tahun 2021 sebanyak 882.591 jiwa. Terjadi penambahan penduduk sebanyak 14.669 jiwa dari periode 2020.

#### 4. KESIMPULAN

Maka dapat ditarik kesimpulan dari hasil analisa dan pembahasan ini bahwa teknik yang diimplementasikan yakni metode regresi linear berganda untuk mengestimasi laju pertumbuhan penduduk mendapatkan hasil yang cukup akurat. Metode ini dapat digunakan untuk kepentingan selanjutnya seperti pembangunan software untuk mempermudah BPS Simalungun dalam menghitung estimasi dan perkiraan lain dengan berlandaskan metode regresi linear berganda. Hasil ini dapat membantu Badan Pusat Statistik (BPS) Simalungun dalam mencari hasil akurat dalam perhitungan laju estimasi. Atribut lainnya seperti jumlah laki-laki dan perempuan dapat ditemukan pola baru untuk estimasi laju pertumbuhan penduduk

#### REFERENCES

- [1] Yuli Mardi, "Jurnal Edik Informatika Data Mining: Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database ( KDD ) . Jurnal Edik Informatika," *JurnalEdikInformatikaPenelitian Bid. Komput. Sains dan Pendidik. Inform.*, vol. 2, pp. 213–219, 2016.
- [2] A. P. Windarto, "Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering," *Techno.COM*, vol. 16, no. 4, pp. 348–357, 2017.
- [3] Z. R. S. Elsi *et al.*, "Utilization of Data Mining Techniques in National Food Security during the Covid-19 Pandemic in Indonesia," *J. Phys. Conf. Ser.*, vol. 1594, no. 1, 2020, doi: 10.1088/1742-6596/1594/1/012007.
- [4] Sudirman, A. P. Windarto, and A. Wanto, "Data mining tools | rapidminer: K-means method on clustering of rice crops by province as efforts to stabilize food crops in Indonesia," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 420, no. 1, 2018, doi: 10.1088/1757-899X/420/1/012089.
- [5] A. Waluyo, H. Jatnika, M. R. S. Permatasari, T. Tuslaela, I. Purnamasari, and A. P. Windarto, "Data Mining Optimization uses C4.5 Classification and Particle Swarm Optimization (PSO) in the location selection of Student Boardinghouses," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 874, no. 1, pp. 1–9, 2020, doi: 10.1088/1757-899X/874/1/012024.
- [6] A. P. Windarto, U. Indriani, M. R. Raharjo, and L. S. Dewi, "Bagian 1: Kombinasi Metode Klastering dan Klasifikasi (Kasus Pandemi Covid-19 di Indonesia)," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 855, 2020, doi: 10.30865/mib.v4i3.2312.
- [7] D. Firdaus, "Penggunaan Data Mining dalam Kegiatan Sistem Pembelajaran Berbantuan Komputer," vol. 6, pp. 91–97, 2017.
- [8] D. H. Ningsih S., "Penerapan Metode Suksesif Interval pada Analisis Regresi," vol. 1, pp. 43–53, 2019.