Penerapan Metode K-Means Dalam Mengelompokkan Banyaknya Desa/ Kelurahan Menurut Keberadaan Permukiman Di Bantaran Sungai Berdasarkan Provinsi

Indah Pratiwi M.S¹, Agus Perdana Windarto², Irfan Sudahri Damanik³
^{1,2,3}STIKOM Tunas Bangsa, Pematangsiantar, Sumatera Utara, Indonesia
 Jln. Sudirman Blok A No. 1-3 Pematangsiantar, Sumatera Utara
¹indahpratiwi1812@gmail.com, ²agus.perdana@amiktunasbangsa.ac.id

Abstract

The research aims to classify the settlements along the river banks by province. To solve this problem, the researchers applied the K-Means Algorithm method. Where the source of research data was collected based on documents explaining the number of villages / sub-districts according to the existence of settlements on the river banks produced by the Central Statistics Agency (BPS). The data used in the study are data from 2014 - 2018 which consists of 34 provinces. The data will be processed by clustering in 2 clusters, namely the settlement level cluster on the high riverbank and the settlement level cluster on the low riverbank. The high cluster consists of 11 data, namely the provinces of Aceh, North Sumatra, Jambi, South Sumatra, West Java, Central Java, East Java, West Kalimantan, Central Kalimantan, South Kalimantan, and South Sulawesi. By conducting the research, it can provide input and as a solution to related parties in charge of dealing with settlement problems along the river banks, especially for the government, in order to get more attention in provinces with high riverbank settlement rates.

Keywords : Data Mining, Riverbank Settlements, Clustering, K-Means

Abstrak

Penelitian bertujuan untuk mengelompokkan permukiman di bantaran sungai berdasarkan provinsi. Untuk menyelesaikan permasalahan ini peneliti menerapkan metode Algoritma K-Means. Dimana sumber data penelitian dikumpulkan berdasarkan dokumen – dokumen keterangan Jumlah Desa/Kelurahan Menurut Keberadaan Permukiman Di Bantaran Sungai yang dihasilkan oleh Badan Pusat Statistik (BPS). Data yang digunakan dalam penelitian adalah data dari tahun 2014 – 2018 yang terdiri dari 34 provinsi. Data akan diolah dengan melakukan clustering dalam 2 cluster yaitu cluster tingkat permukiman di bantaran sungai tinggi dan cluster tingkat permukiman di bantaran sungai rendah. Cluster tinggi berjumlah 11 data yaitu provinsi Aceh, Sumatera Utara, Jambi, Sumatera Selatan, Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, dan Sulawesi Selatan. Dengan dilakukannya penelitian dapat memberi masukan dan sebagai solusi kepada pihak terkait yang bertugas menangani permasalahan permukiman di bantaran sungai khususnya terhadap pemerintah, agar mendapat perhatian lebih pada provinsi dengan tingkat permukiman di bantaran sungai tinggi.

Kata Kunci : Data Mining, Permukiman di Bantaran Sungai, Clustering, K-Means

1. Pendahuluan

Permukiman di Indonesia masih banyak yang lingkungannya tidak layak dijadikan permukiman seperti permukiman yang berada di bantaran sungai. Bantaran sungai merupakan ruang antara tepi palung sungai dan kaki tanggul sebelah dalam yang terletak di kiri dan di kanan palung sungai [1]. Permukiman berada di bantaran sungai dikarenakan lahan untuk permukiman yang semakin berkurang dan karena tingginya harga lahan untuk membeli perumahan. Karena

keterbatasan lahan maka dibuatlah pemadatan bangunan yang menjadi sebab utama permukiman dibangun disekitaran bantaran sungai dengan kualitas lingkungan rendah.

Semakin banyak masyarakat yang mendirikan tempat tinggal di bantaran sungai maka akan menimbulkan berbagai masalah seperti : permukiman yang rawan terkena banjir dan timbulnya berbagai macam penyakit diantaranya penyakit kulit, diare, dan cacingan yang diakibatkan dari pemanfaatan air sungai untuk kehidupan sehari – hari. Sehingga masyarakatnya kesulitan untuk hidup sehat. Hal ini sangat membahayakan masyarakat yang tinggal di sekitar bantaran sungai. Maka perlu dilakukan pengelompokan untuk mengetahui provinsi yang memiliki tingkat permukiman di bantaran sungainya masih tinggi.

Oleh sebab itu solusi yang dapat digunakan untuk memecahkan permasalahan tersebut yaitu penulis menggunakan metode *Data Mining* algoritma *K-means*. *K-means* merupakan suatu algoritma yang digunakan dalam pengelompokan secara pertisi yang memisahkan data ke dalam kelompok yang berbeda – beda. Algoritma ini mampu meminimalkan jarak antara data ke *clusternya* [2]. Algoritma *K-means* penggunaannya sudah diterapkan pada beberapa penelitian yang sudah pernah dilakukan seperti penelitian yang dilakukan oleh [3] berjudul “Implementasi *Data Mining* dalam Mengelompokkan Rumah Tangga Kumuh di Perkotaan Berdasarkan Provinsi Menggunakan Algoritma *K-Means*”. Berdasarkan penelitian yang dilakukan, bisa diambil kesimpulan bahwa algoritma *K-means* telah berhasil mengelompokkan rumah tangga kumuh yang ada di provinsi Indonesia dengan ditentukannya nilai *centroid* dalam 3 *cluster* yaitu *cluster* tinggi (C1), *cluster* sedang (C2), dan *cluster* rendah (C3).

Diharapkan dengan menggunakan metode ini dapat dihasilkan *clustering* yang terbukti akurat dalam kasus mengelompokkan permukiman di bantaran sungai berdasarkan provinsi yang ada di Indonesia. Maka akan diketahui provinsi mana saja di Indonesia yang memiliki permukiman di sekitar bantaran sungai tinggi dan rendah.

2. Metodologi Penelitian

2.1. Data Mining

Data mining merupakan metode yang digunakan untuk pengolahan data guna menemukan pola tersembunyi dari data yang akan diolah. *Data mining* juga merupakan rangkaian kegiatan untuk menemukan suatu pola yang menarik dari data dalam jumlah yang besar, kemudian data – data tersebut dapat disimpan dalam sebuah *data warehouse*, *database*, atau penyimpanan informasi [4].

Menurut [5] tahapan pada proses *data mining* dimulai dengan seleksi data dari data sumber ke data target, tahap *pre-processing* untuk memperbaiki kualitas data, transformasi, *data mining* serta tahap interpretasi dan evaluasi yang kemudian menghasilkan *output* berupa pengetahuan baru dan diharapkan dapat memberikan kontribusi yang lebih baik.

2.2. Algoritma *K-means*

Algoritma *K-means* merupakan salah satu algoritma dengan *partitional*, karena *K-means* didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai *centroid* awalnya. Algoritma *K-means* menggunakan proses secara berulang-ulang untuk mendapatkan basis data *cluster* [6]. Karakteristik dari algoritma *K-means* salah satunya adalah sangat sensitif dalam penentuan titik pusat awal *cluster* karena *K-means* membangkitkan titik pusat *cluster* awal secara random [7].

Langkah – langkah melakukan *clustering* dengan menggunakan algoritma *K-means* adalah sebagai berikut [7]:

- a) Tentukan nilai k sebagai jumlah *cluster* yang ingin dibentuk.

- b) Bangkitkan k *centroid* (titik pusat *cluster*) awal secara random/acak, kemudian untuk menghitung *centroid cluster* ke- i berikutnya, digunakan rumus sebagai berikut :

$$v = \sqrt{\frac{\sum_{i=1}^n x_i}{n}} \quad (1)$$

Dimana :

i : 1,2,3,...n

v : *centroid* pada *cluster*

x_i : objek ke- i

n : banyaknya jumlah objek yang menjadi anggota *cluster*

- c) Hitung jarak setiap data ke masing-masing *centroid* menggunakan rumus.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

Ket :

dij : Jarak objek antara objek x dan y

n : Jumlah Atribut

x_i : Objek Data

y_i : Data *Cluster*

- d) Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroidnya* (C).
e) Kembali ke langkah 3 jika posisi *centroid* baru dengan *centroid* lama tidak sama.

3. Hasil dan Pembahasan

Data yang digunakan dalam penelitian adalah data banyaknya desa/kelurahan menurut keberadaan permukiman di bantaran sungai (desa) pada tahun 2014 dan 2018. Dalam penelitian dikelompokkan menjadi 2 bagian yaitu permukiman di bantaran sungai tinggi dan rendah. Kumpulan data yang penulis peroleh digunakan sebagai data masukan dalam membuat model aturan menggunakan algoritma *K-means* dan menggunakan *software Rapidminer*.

Tabel 1. Data Permukiman Di Bantaran Sungai

No	Provinsi	2014 (x)	2018 (y)
1	Aceh	913	540
2	Sumatera Utara	1019	730
3	Sumatera Barat	491	380
4	Riau	639	412
5	Jambi	843	492
6	Sumatera Selatan	1285	700
7	Bengkulu	437	99
8	Lampung	473	353
9	Kep. Bangka Belitung	33	31
10	Kep. Riau	77	62
11	Dki Jakarta	127	107
12	Jawa Barat	1949	1419
13	Jawa Tengah	1876	916
14	Di Yogyakarta	134	54
15	Jawa Timur	1433	862
16	Banten	516	575
17	Bali	216	122
18	Nusa Tenggara Barat	493	134
19	Nusa Tenggara Timur	282	305
20	Kalimantan Barat	900	718

No	Provinsi	2014 (x)	2018 (y)
21	Kalimantan Tengah	1012	910
22	Kalimantan Selatan	1242	1039
23	Kalimantan Timur	467	304
24	Kalimantan Utara	143	54
25	Sulawesi Utara	664	387
26	Sulawesi Tengah	638	454
27	Sulawesi Selatan	952	670
28	Sulawesi Tenggara	438	153
29	Gorontalo	225	110
30	Sulawesi Barat	286	197
31	Maluku	185	146
32	Maluku Utara	110	174
33	Papua Barat	201	150
34	Papua	366	319

a) Menentukan Nilai k Jumlah *Cluster*

Jumlah *cluster* permukiman di bantaran sungai sebanyak 2 *cluster*. *Cluster* yang dibentuk yaitu *cluster* tinggi (C1) dan *cluster* rendah (C2).

b) Menentukan Nilai *Centroid* (Pusat *Cluster*)

Penentuan pusat *cluster* awal ditentukan secara *random* yang diambil dari data yang ada dalam *range*. Adapun nilai untuk *cluster* tinggi (*cluster* 1) diambil dari nilai tertinggi yang terdapat pada tabel 1 dan nilai untuk *cluster* rendah (*cluster* 2) diambil dari nilai terendah yang terdapat pada tabel 1.

Tabel 2. *Centroid* Data Awal Iterasi 1

<i>Cluster</i>	Nilai	
	X	Y
C1	1949	1419
C2	33	31

c) Menghitung Jarak Setiap Data Terhadap *Centroid* (Pusat *Cluster*)

Setelah data nilai pusat *cluster* awal ditentukan, maka langkah selanjutnya adalah menghitung jarak masing-masing data terhadap pusat *cluster*. Untuk menghitung jarak setiap data permukiman di bantaran sungai terhadap pusat *cluster* dapat menggunakan rumus yang perhitungannya dapat kita lihat sebagai berikut :

Dilakukan perhitungan jarak terhadap data permukiman di bantaran sungai dengan titik pusat (*centroid*) pada *cluster* pertama.

$$D_{(1.1)} = \sqrt{((913-1949)^2 + (540-1419)^2)} = 1358,65$$

$$D_{(1.2)} = \sqrt{((1019-1949)^2 + (730-1419)^2)} = 1157,42$$

$$D_{(1.3)} = \sqrt{((491-1949)^2 + (380-1419)^2)} = 1790,33$$

$$D_{(1.4)} = \sqrt{((639-1949)^2 + (412-1419)^2)} = 1652,32$$

$$D_{(1.5)} = \sqrt{((843-1949)^2 + (492-1419)^2)} = 1443,11$$

Perhitungan jarak permukiman di bantaran sungai pertama dengan *centroid cluster* kedua, seperti berikut :

$$D_{(2.1)} = \sqrt{((913-33)^2 + (540-31)^2)} = 1016,60$$

$$D_{(2,2)} = \sqrt{((1019-33)^2 + (730-31)^2)} = 1208,63$$

$$D_{(2,3)} = \sqrt{((491-33)^2 + (380-31)^2)} = 575,82$$

$$D_{(2,4)} = \sqrt{((639-33)^2 + (412-31)^2)} = 715,82$$

$$D_{(2,5)} = \sqrt{((843-33)^2 + (492-31)^2)} = 932,00$$

Berikut tabel 3 hasil perhitungan jarak data dengan titik pusat *cluster* pada iterasi 1 menggunakan *Eulidean Distance*.

Tabel 3. Jarak *Centroid* Iterasi 1

No	Provinsi	Tinggi	Rendah	Jarak Terpendek	C1	C2
1	Aceh	1358,65	1016,60	1016,60		1
2	Sumatera Utara	1157,42	1208,63	1157,42	1	
3	Sumatera Barat	1790,33	575,82	575,82		1
4	Riau	1652,32	715,82	715,82		1
5	Jambi	1443,11	932,00	932,00		1
6	Sumatera Selatan	978,70	1419,53	978,70	1	
7	Bengkulu	2007,12	409,68	409,68		1
8	Lampung	1820,70	545,24	545,24		1
9	Kep. Bangka Belitung	2365,92	0,00	0,00		1
10	Kep. Riau	2312,11	53,82	53,82		1
11	Dki Jakarta	2245,22	120,88	120,88		1
12	Jawa Barat	0,00	2365,92	0,00	1	
13	Jawa Tengah	508,27	2044,47	508,27	1	
14	Di Yogyakarta	2271,00	103,59	103,59		1
15	Jawa Timur	759,28	1628,05	759,28	1	
16	Banten	1663,08	727,48	727,48		1
17	Bali	2164,60	204,38	204,38		1
18	Nusa Tenggara Barat	1941,95	471,39	471,39		1
19	Nusa Tenggara Timur	2004,97	370,24	370,24		1
20	Kalimantan Barat	1261,67	1106,19	1106,19		1
21	Kalimantan Tengah	1066,33	1315,71	1066,33	1	
22	Kalimantan Selatan	802,65	1574,09	802,65	1	
23	Kalimantan Timur	1854,60	512,72	512,72		1
24	Kalimantan Utara	2263,82	112,38	112,38		1
25	Sulawesi Utara	1648,10	724,50	724,50		1
26	Sulawesi Tengah	1627,87	738,21	738,21		1
27	Sulawesi Selatan	1247,00	1119,32	1119,32		1
28	Sulawesi Tenggara	1971,26	422,98	422,98		1
29	Gorontalo	2164,64	207,62	207,62		1
30	Sulawesi Barat	2063,70	302,60	302,60		1
31	Maluku	2175,37	190,60	190,60		1
32	Maluku Utara	2220,80	162,41	162,41		1
33	Papua Barat	2160,06	205,88	205,88		1
34	Papua	1927,66	440,26	440,26		1

d) Menentukan Posisi *Cluster* atau Pengelompokan

Dalam menentukan posisi *cluster* masing-masing data permukiman di bantaran sungai berdasarkan jarak minimum data terhadap pusat *cluster*. Data yang

memiliki jarak terkecil dengan *centroid* akan menjadi anggota pada kelompok tersebut. Posisi *cluster* dapat dilihat pada tabel 3 posisi data dengan tiap *cluster* pada iterasi 1 dengan menggunakan bantuan angka 1 sebagai tanda yang mengatakan data menjadi anggota pada *cluster*.

Selanjutnya dalam metode *K-means*, perhitungan berhenti apabila *cluster* pada iterasi yang dihasilkan sama pada iterasi sebelumnya. Maka selanjutnya mencari *cluster* pada iterasi selanjutnya sampai nilai iterasinya sama.

- e) Menghitung *Centroid* Baru Menggunakan Hasil Dari Setiap Anggota Pada Masing-Masing *Cluster*

Untuk mencari nilai *centroid* selanjutnya dengan menggunakan *centroid* baru pada iterasi 1 dengan menjumlahkan nilai sesuai yang tertera pada *cluster* yang terdapat pada tabel 3. Adapun *centroid* baru untuk mencari *cluster* selanjutnya adalah dengan menjumlahkan nilai yang terpilih pada *cluster* tersebut kemudian membagikannya sebanyak jumlah nilai.

Contoh perhitungan titik pusat baru pada *cluster* x dan y adalah sebagai berikut :

$$C_1 x = \frac{1019+1285+1949+1876+1433+1012+1242}{7} = 1402,29$$

$$C_1 y = \frac{730+700+1419+916+862+910+1039}{7} = 939,43$$

$$C_2 x = \frac{913+491+639+843+437+473+33+77+127+134+516+216+493+282+900+467+143+664+638+952+438+225+286+185+110+201+366}{27} = 416,63$$

$$C_2 y = \frac{540+380+412+492+99+353+31+62+107+54+575+122+134+305+718+304+54+387+454+670+153+110+197+146+174+150+319}{27} = 277,85$$

Maka, *centroid* data baru iterasi 1 adalah sebagai berikut :

Tabel 4. *Centroid* Data Baru Iterasi 2

Cluster	Nilai	
	X	Y
C1	1402,29	939,43
C2	416,63	277,85

- f) Selanjutnya dilakukan kembali langkah ke 4 sampai 6.

Jika nilai *centroid* hasil iterasi dengan nilai *centroid* sebelumnya bernilai sama atau nilai *centroid* sudah optimal serta posisi *cluster* data permukiman di bantaran sungai tidak mengalami perubahan lagi maka proses iterasi berhenti. Namun jika nilai *centroid* tidak sama atau belum optimal serta posisi data masih berubah maka proses iterasi berlanjut pada iterasi berikutnya. Pada penelitian ini proses perhitungan dilakukan hingga iterasi 5. Berikut adalah hasil perhitungan jarak dengan titik pusat cluster pada iterasi 5.

Tabel 5. Jarak *Centroid* Iterasi 5

No	Provinsi	Tinggi	Rendah	Jarak Terpendek	C1	C2
1	Aceh	414,31	662,64	414,31	1	
2	Sumatera Utara	219,68	854,87	219,68	1	
3	Sumatera Barat	850,68	224,74	224,74		1
4	Riau	708,99	361,40	361,40		1
5	Jambi	498,56	578,24	498,56	1	
6	Sumatera Selatan	134,38	1066,43	134,38	1	
7	Bengkulu	1063,18	160,79	160,79		1
8	Lampung	880,12	193,02	193,02		1
9	Kep. Bangka Belitung	1424,40	354,42	354,42		1
10	Kep. Riau	1370,60	300,67	300,67		1

No	Provinsi	Tinggi	Rendah	Jarak Terpendek	C1	C2
11	Dki Jakarta	1304,11	234,73	234,73		1
12	Jawa Barat	944,63	2012,29	944,63	1	
13	Jawa Tengah	662,95	1693,03	662,95	1	
14	Di Yogyakarta	1328,01	259,16	259,16		1
15	Jawa Timur	217,18	1273,84	217,18	1	
16	Banten	745,04	398,90	398,90		1
17	Bali	1221,85	152,64	152,64		1
18	Nusa Tenggara Barat	998,33	182,79	182,79		1
19	Nusa Tenggara Timur	1069,35	97,90	97,90		1
20	Kalimantan Barat	335,55	754,61	335,55	1	
21	Kalimantan Tengah	227,84	967,93	227,84	1	
22	Kalimantan Selatan	222,24	1223,48	222,24	1	
23	Kalimantan Timur	911,90	158,31	158,31		1
24	Kalimantan Utara	1320,66	252,34	252,34		1
25	Sulawesi Utara	703,67	371,01	371,01		1
26	Sulawesi Tengah	686,67	384,46	384,46		1
27	Sulawesi Selatan	306,38	765,36	306,38	1	
28	Sulawesi Tenggara	1026,68	125,73	125,73		1
29	Gorontalo	1221,37	154,29	154,29		1
30	Sulawesi Barat	1121,81	52,06	52,06		1
31	Maluku	1234,23	165,20	165,20		1
32	Maluku Utara	1283,51	227,12	227,12		1
33	Papua Barat	1218,64	149,17	149,17		1
34	Papua	989,32	103,70	103,70		1

Perhitungan manual pada data permukiman di bantaran sungai di atas didapatkan hasil akhir yang dimana pada iterasi 5 pengelompokkan data yang dilakukan terhadap 2 *cluster* dengan iterasi 4 didapatkan hasil yang sama. Hasil dari kedua iterasi tersebut bernilai C1 = 11 dan C2 = 23 pada posisi data tiap *cluster* x dan y. Sehingga posisi *cluster* pada data tersebut tidak mengalami perubahan lagi maka proses iterasi berhenti.

Berdasarkan dari penjelasan di atas mengenai tahap-tahap penyelesaian algoritma k-means secara manual, maka berikut akan dibahas mengenai keterkaitan dari hasil yang didapat antara perhitungan manual algoritma dengan hasil yang ditampilkan oleh *tools Rapidminer* dan dapat disimpulkan bahwa hasil perhitungan manual dari metode *K-means* dan dari aplikasi yaitu *software Microsoft Excel 2010* hasilnya sama dan dapat di *input* ke dalam aplikasi *Rapidminer*.

Row No.	A	cluster	2014.0	2018.0
9	Kep. Bangka	cluster_1	33	31
10	Kep. Riau	cluster_1	77	62
11	Dki Jakarta	cluster_1	127	107
14	Di Yogyakarta	cluster_1	134	54
16	Banten	cluster_1	516	575
17	Bali	cluster_1	216	122
18	Nusa Tengg	cluster_1	493	134
19	Nusa Tengg	cluster_1	282	305
23	Kalimantan	cluster_1	467	304
24	Kalimantan	cluster_1	143	54
25	Sulawesi Ut	cluster_1	664	387
26	Sulawesi Te	cluster_1	638	454
28	Sulawesi Te	cluster_1	438	153
29	Gorontalo	cluster_1	225	110
30	Sulawesi Ba	cluster_1	286	197
31	Maluku	cluster_1	185	146
32	Maluku Utara	cluster_1	110	174
33	Papua Barat	cluster_1	201	150
34	Papua	cluster_1	366	319

Gambar 1. Tampilan Data Perhitungan *Tools Rapidminer*

Berdasarkan gambar 1 dapat dilihat bahwa posisi *cluster* masing-masing data permukiman di bantaran sungai dan nilai *cluster* hasil iterasi 5 maka dapat disimpulkan bahwa :

- a) *Cluster* tinggi (C1) dengan jumlah data permukiman di bantaran sungai sebanyak 11 provinsi yaitu : Aceh, Sumatera Utara, Jambi, Sumatera Selatan, Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, dan Sulawesi Selatan.
- b) *Cluster* rendah (C2) dengan jumlah data permukiman di bantaran sungai sebanyak 23 provinsi yaitu : Sumatera Barat, Riau, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua.

4. Kesimpulan

Berdasarkan pembahasan sebelumnya dapat disimpulkan bahwa :

- a) *Data Mining* dengan algoritma *K-means* dapat diterapkan untuk mengelompokkan banyaknya desa/kelurahan menurut keberadaan permukiman di bantaran sungai berdasarkan provinsi. Sumber data yang digunakan pada penelitian ini adalah data yang diperoleh dari BPS. Jumlah data yang data uji sebanyak 34 provinsi dengan menggunakan dua *cluster*, yaitu untuk *cluster* tinggi berjumlah 11 data yaitu provinsi Aceh, Sumatera Utara, Jambi, Sumatera Selatan, Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, dan Sulawesi Selatan. *Cluster* rendah berjumlah 23 provinsi.
- b) Hasil yang diperoleh dari metode *K-means* dapat diterapkan ke dalam *rapidminer* dengan nilai validasi yang sama dan dapat menjadi masukan kepada pemerintah agar dilakukan penanganan lebih dan antisipasi untuk membantu memperbaiki permukiman di bantaran sungai

Daftar Pustaka

- [1] B. T. Cesarin And C. Ginting, "Persepsi Masyarakat Terhadap Permukiman Bantaran Sungai," No. 1, Pp. 117–122, 2015.
- [2] R. W. Sari, A. Wanto, And A. P. Windarto, "Implementasi Rapidminer Dengan Metode K-Means (Study Kasus : Imunisasi Campak Pada Balita Berdasarkan Provinsi)," Vol. 2, Pp. 224–230, 2018.
- [3] R. A. Margolang, S. R. Andani, And M. R. Lubis, "Implementasi Data Mining Dalam Mengelompokkan Rumah Tangga Kumuh Di Perkotaan Berdasarkan Provinsi Menggunakan," No. September, Pp. 602–609, 2019.
- [4] M. G. Sadewo, A. P. Windarto, And S. R. Andani, "Pemanfaatan Algoritma Clustering Dalam Mengelompokkan Jumlah Desa / Kelurahan Yang Memiliki Sarana Kesehatan," Vol. I, Pp. 124–131, 2017.
- [5] K. Fatmawati And A. P. Windarto, "Data Mining : Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi," Vol. 3, No. 2, Pp. 173–178, 2018.
- [6] N. Putu, E. Merliana, P. Studi, M. Teknik, F. T. Industri, And U. A. Jaya, "Analisa Penentuan Jumlah Cluster Terbaik Pada Metode K-Means," Pp. 978–979.
- [7] N. V. Waworuntu And M. F. Amin, "Penerapan Metode K-Means Pemetaan Calon," Vol. 05, No. 02, Pp. 190–200, 2018.