

BAB II

TINJAUAN PUSTAKA

2.1 Teori Umum

2.1.1 Data

Data adalah kumpulan fakta maupun informasi yang secara acak belum memiliki makna, baik berupa angka , simbol atau kata-kata yang diperoleh dari hasil pengamatan tertentu. Seperti dikutip oleh Tom Stonier dalam jurnal Wolski & Gomolińska (2020) Data adalah serangkaian fakta dan observasi yang tidak berhubungan. Jenis data ada 2 yaitu data kualitatif dan data kuantitatif. Data kualitatif biasanya berbentuk angka , simbol, maupun gambar.

2.1.3 Informasi

Informasi merupakan data yang dimasukkan ke dalam struktur yang sesuai untuk penerima manfaat dan membantu secara dinamis (Rudi Hermawan, Arief Hidayat, 2016). Syarat informasi mengacu pada beberapa semantik (makna) data, yang pada gilirannya dapat dinyatakan secara eksplisit (Wolski & Gomolińska, 2020). Agar dapat memiliki makna, informasi harus memiliki bagian penting meliputi ciri, dapat dipastikan kebenarannya, relevan dengan kondisi yang ada, serta mempunyai hubungan dengan waktu yang diambil. Informasi juga dapat dikatakan berkualitas apabila dapat diverifikasi.

2.1.2 Data Science

Information science was characterized as "the study of managing information, whenever they have been set up, while the connection of the

information to what they address is assigned to different fields and sciences (Cao, 2017). Sedangkakan menurut Vashant Dhar *Information science has been characterized as a "idea to bind together insights, information examination, AI and their connected strategies" to "comprehend and break down genuine wonders" with information* (Dhar, 2013). Menurut (Nainggolan, 2017) tahapan dalam *data science* serupa dengan 3 tahapan *Knowledge Discovery in Database* (KDD):

1. Pra-pemrosesan Data untuk menganalisis kumpulan data multivariat sebelum penambangan data.
2. Data Mining yang umumnya melibatkan enam kelompok tugas:
 - a. Deteksi anomali (deteksi penyimpangan)
 - b. Mengidentifikasi catatan data yang tidak biasa, data menarik atau kesalahan data yang memerlukan penyelidikan lebih lanjut.
 - c. Pembelajaran aturan asosiasi (pemodelan dependensi)
 - d. Menemukan hubungan antar variabel
 - e. *Clustering* (pengelompokan)
 - f. Klasifikasi
 - g. Regresi
 - h. Ringkasan
3. Validasi Hasil adalah terstruktur memahami dan memberi makna pada hasil dari jenis prediksi yang sama atau temukan tipe pola umum yang sama terlepas dari data uji yang disediakan.

2.1.4 Data Mining Software

Software data mining mengacu pada perangkat lunak yang memungkinkan perusahaan dan pengguna lain untuk mengekstrak data yang dapat digunakan dari sekumpulan besar data mentah untuk menemukan korelasi, pola, dan anomali. Hasil dari proses data mining membantu perusahaan memprediksi hasil (Research, 2020). Ada berbagai software data mining yang banyak digunakan dalam penelitian seperti WEKA (*Waikato Environment for Knowledge Analysis*), Rapid Miner Studio, Rattle, Orange, dan Knime . Umumnya software data mining ini digunakan dalam berbagai kepentingan dengan berbagai fitur yang telah disediakan dalam *software* seperti Kmeans, Hierarchical dan berbagai algoritma yang lain.

2.1.5 Pandemi

Menurut Kamus Besar Bahasa Indonesia pandemi merupakan wabah yang berjangkit serempak di mana-mana, meliputi daerah geografi yang luas (KBBI Daring, n.d.). Maka dapat dikatakan bahwa penyakit ini merupakan masalah bersama bagi seluruh dunia. Contoh penyakit yang tergolong pandemi adalah *Influenza* dan COVID-19.

2.1.5.1 Sejarah Pandemi

Dalam buku Waspada Penyakit Flu Burung Abdullah Pantan yang dikutip oleh de Rosari & Saepudin (2013) Sejarah mencatat bahwa menjelang awal abad kedua puluh, pandemi flu utama terjadi, pada tahun 1918 pandemi utama yang disebabkan oleh infeksi flu H1N1 atau disebut juga influenza Spanyol yang terjadi di tiga wilayah berbeda yang berjauhan, khususnya Brest di Prancis, Boston di Amerika Serikat dan Freetown di Sierra Leone. . Flu Spanyol memiliki tingkat

kerusakan yang sangat signifikan, menyerang orang dewasa 20-40 tahun, dan membuat 180-360 juta orang meninggal dunia. Karena influenza Spanyol pada tahun 1957-1958, influenza Asia subtipe H2N2 pertama kali ditemukan di Cina dan selama dua tahun penyebarannya telah menewaskan 70.000 orang Amerika. Pada tahun 1968-1969 terjadi influenza H3N2 Hong Kong yang membuat 2-7,5 juta orang kehilangan nyawa, dan pada tahun 1977 terjadi influenza Rusia..

2.1.5.2 Jenis Pandemi

Selain pandemi covid-19 telah banyak pandemi yang pernah terjadi di dunia seperti yang dikutip oleh (B2P2VRP Salatiga, 2020) berikut jenis pandemi yang pernah terjadi:

1. Pes
2. Kolera
3. Flu Spanyol
4. Flu Asia
5. Flu Hongkong
6. HIV/AIDS
7. SARS (*Severe Acute Respiratory Syndrome*)
8. Flu Babi
9. Ebola

2.1.6 Bantuan Sosial

Sesuai dengan Peraturan Menteri Keuangan Republik Indonesia Nomor 254/Pmk.05/2015 tentang Belanja Bantuan Sosial Pada Kementerian Negara/Lembaga, Bantuan Sosial Tahun 2015 adalah konsumsi sebagai transfer

uang tunai, produk atau administrasi yang disediakan oleh Pemerintah kepada masyarakat miskin atau tidak layak untuk melindungi daerah dari kemungkinan bahaya sosial, memperluas batas keuangan dan bantuan tambahan pemerintah daerah. Ada beberapa jenis bantuan sosial yang diberikan otoritas publik ke daerah, seperti Program Indonesia Pintar (PIP), Program Jaminan Kesehatan Nasional (JKN-KIS), Program Keluarga Harapan (PKH), dan Rastra/ Bantuan Sosial/Bantuan Pangan Non Tunai. (RI, 2018).

2.1.7 Penduduk Miskin

BPS menggunakan gagasan kapasitas untuk mengatasi isu-isu esensial (*fundamental need approach*). Dengan metodologi ini, kemiskinan dipandang sebagai ketidakberdayaan finansial untuk memenuhi kebutuhan pangan dan non-pangan yang esensial yang diperkirakan dari sisi penggunaan. (STATISTIK, 2021). Penduduk dikatakan miskin jika memenuhi beberapa kriteria menurut Badan Pusat Statistik (BATAM, 2014).

1. Luas lantai bangunan rumah kurang dari 8m² per orang
2. Jenis lantai rumah terbuat dari tanah/bambu/kayu murahan
3. Jenis dinding rumah dari bambu/ rumbia/ kayu berkualitas rendah/tembok tanpa diplester.
4. Tidak mempunyai fasilitas buang air besar/ bersama-sama dengan rumah tangga lain.
5. Sumber penerangan rumah tangga tidak menggunakan listrik.
6. Sumber air minum berasal dari sumur/ mata air tidak terlindung/ sungai/ air hujan.

7. Bahan bakar untuk memasak sehari-hari adalah kayu bakar/ arang/ minyak tanah
8. Hanya mengkonsumsi daging/ susu/ ayam dalam satu kali seminggu.
9. Hanya membeli satu stel pakaian baru dalam setahun
10. Hanya mampu makan sebanyak satu/ dua kali dalam sehari
11. Tidak mampu membayar biaya pengobatan di puskesmas/ poliklinik
12. Jenis pendapatan kepala keluarga adalah: peternak dengan luas lahan 500m², buruh tani, pemancing, buruh bangunan, buruh tani dan tambahan jabatan lain dengan upah di bawah Rp. 600.000,- setiap bulan
13. Pendidikan tertinggi kepala keluarga : tidak sekolah/ tidak tamat SD/ tamat SD.
14. Tidak mempunyai tabungan/ barang yang dapat dijual dengan minimal Rp. 500.000,- seperti sepeda motor kredit/ non kredit, emas, ternak, kapal motor, atau barang modal lainnya.

2.1.8 Tenaga Kerja

Berdasarkan Undang-Undang Republik Indonesia Nomor 13 Tahun 2003 Tentang Ketenagakerjaan Tenaga kerja adalah setiap orang yang dapat mengelola pekerjaan untuk menghasilkan produk atau potensi manfaat baik untuk mengatasi masalah mereka sendiri maupun untuk daerah. Pada akhirnya, dapat dikatakan bahwa setiap orang yang tidak bekerja adalah orang yang tidak memiliki kemampuan untuk membuat produk dan jasa untuk memenuhi kebutuhan hidup..

2.2 Teori Khusus

2.2.1 Bantuan Sosial Tunai (BST)

Bantuan Sosial Uang (BST) adalah bantuan berupa uang tunai yang diberikan kepada keluarga miskin, tertindas, serta lemah yang dipengaruhi oleh merebaknya Corona Virus Disease 2019 (COVID-19). Besaran Bantuan Sosial Tunai sebesar Rp. 600.000,-/keluarga/bulan. Bantuan Sosial Tunai ini merupakan program jaringan jaminan sosial Kementerian Sosial yang diperuntukkan bagi keluarga miskin dan lemah terdampak Covid-19. Program ini merupakan tugas bantuan yang tidak biasa kepada Presiden. Bantuan sosial untuk daerah di luar Jabodetabek diberikan dalam bentuk tunai, sedangkan untuk daerah Jabodetabek diberikan sebagai kebutuhan pokok. Pemberian bantuan BST, kecuali penerima Program Keluarga Harapan (PKH), Kartu Sembako, dan Kartu Prakerja (STATISTIK, 2021).

2.2.3 Bantuan Pangan Non Tunai (BPNT)/Program Sembako

Bantuan Pangan Non Tunai (BPNT)/Program Sembako adalah bantuan sosial pangan dengan jenis nontunai dari badan publik yang diberikan kepada Keluarga Penerima Manfaat (KPM) secara konsisten melalui komponen finansial. KPM akan mendapatkan bantuan non tunai berupa kupon elektronik (e-voucher) dari Bank Penyalur. Besaran Bantuan Pangan Non Tunai sebesar Rp 110.000 per KPM setiap bulan untuk BPNT. Sedangkan ukuran program sembako periode Januari-Februari adalah Rp. 150.000,- namun sejak periode Maret-Agustus 2020 menjadi Rp. 200.000,- . Bantuan tersebut tidak dapat diambil dengan uang sungguhan dan jika bantuan tersebut tidak habis dalam bulan tersebut, maka nilai

bantuan tersebut masih disimpan dan dikumpulkan. KPM dapat memanfaatkan e-voucher untuk membeli beras dan bahan makanan lainnya seperti telur, sesuai jumlah dan kualitas yang ideal di e-warong (STATISTIK, 2021)

2.2.4 Program Keluarga Harapan (PKH)

Program Keluarga Harapan (PKH) adalah program bantuan sosial terbatas (bansos) yang diberikan kepada keluarga rentan sesuai aturan yang ditetapkan. Sasaran PKH adalah keluarga miskin dan lemah yang terdaftar dalam Data Kesejahteraan Sosial Terpadu (DTKS) dan memenuhi segmen sejahtera, sekolah, dan tambahan bantuan pemerintah. (STATISTIK, 2021).

2.2.5 Coronavirus

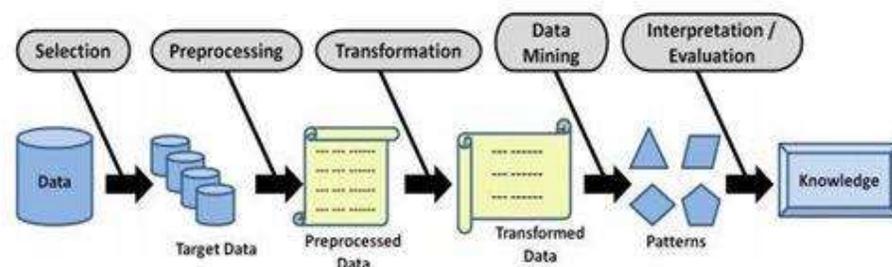
Covid adalah kumpulan infeksi yang dapat menyebabkan penyakit pada makhluk atau manusia. Beberapa jenis Covid diketahui menyebabkan penyakit saluran pernapasan pada orang-orang mulai dari batuk dan pilek hingga penyakit yang lebih serius seperti Middle East Respiratory Syndrome (MERS) dan Severe Acute Respiratory Syndrome (SARS).(WHO, 2021). *Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)* adalah infeksi lain yang mencemari sistem pernapasan individu yang tercemar, infeksi ini sering disebut sebagai Covid-19(Lai et al., 2020). Penyebaran infeksi ini sangat cepat, mengenai cara penyebarannya, terutama melalui percikan ludah , melalui udara, permukaan yang tercemar, melalui kotoran manusia. Selain itu, ada tempat-tempat yang rentan terhadap penyebaran infeksi ini, khususnya tempat-tempat padat, tempat terbatas dan kamar terikat dan tertutup. (Sari, 2021).

2.2.6 Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) diartikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data (Mardalius, 2018) Hasil keluaran yang diperoleh dapat dimanfaatkan sebagai (*knowledge base*) yang dapat dijadikan sebagai dasar pengambilan keputusan (Gustientiedina et al., 2019).

Menurut Elisa (2017) *Knowledge Discovery in Databases* (KDD) adalah cara yang menggabungkan pengumpulan, pemanfaatan informasi yang direkam untuk menemukan normalitas, contoh atau koneksi dalam informasi besar dan asosiasi dengan prosedur rekonsiliasi dan pengungkapan logis, terjemahan dan persepsi contoh dalam berbagai kumpulan informasi..

Secara lebih detail, proses *Knowledge Discovery in Database* (KDD) seperti gambar di bawah ini.



Sumber: (Gustientiedina et al., 2019)

Gambar 2. 1 Proses Knowledge Discovery in Database (KDD)

1. Selection

Selection digunakan untuk mengambil faktor-faktor yang akan diambil agar tidak terjadi kemiripan dan redundansi dalam penanganan data.

2. *Processing*

Terdapat dua tahapan dalam *processing* yaitu:

Data *cleaning* untuk membersihkan data yang tidak digunakan

Data *integration* dilakukan pada kredit yang membedakan substansi baru.

3. *Transformation*

Merubah data sesuai format ekstensi yang sesuai dengan pengolahan data mining.

4. *Data mining*

Proses utama pada metode untuk mendapatkan pengetahuan baru dari data yang di proses.

5. *Evaluation/Interpretation*

Mengidentifikasi pola-pola yang menarik kedalam *knowledge base* yang di identifikasi.

6. *Knowledge*

Pengetahuan yang diperoleh akan diperkenalkan. Pada tahap ini informasi baru yang dibuat dapat sampai pada setiap individu yang akan digunakan sebagai semacam perspektif untuk pengambilan keputusan.

2.2.6.1 Data Mining

Data mining adalah prosedur untuk memperoleh data penting yang disimpan atau tersimpan dalam kumpulan data yang sangat besar atau dalam tumpukan informasi, sehingga ditemukan contoh-contoh menarik yang sebelumnya tidak jelas (Azwanti, 2019). Penambangan data juga merupakan hasil dari mengekstrak

data dalam kumpulan data lingkup besar dan ukuran urutan robotisasi kasus tergantung pada desain informasi yang didapat dari indeks informasi.(Aras & Sardjono, 2016). Penambangan data bertujuan untuk menggali informasi dan pengetahuan yang dapat dimanfaatkan oleh pekerja informasi yang diidentifikasi dengan pemikiran dinamis dan kritis (Waworuntu & Amin, 2018).

Menurut Handoko (2016) Penambangan data adalah siklus yang menggunakan faktual, numerik, penalaran buatan manusia, dan prosedur untuk mengeluarkan dan mengenali data berharga dan informasi terkait dari kumpulan data yang sangat besar. *There are several techniques that data mining has on the task that can be done* (Ardiada et al., 2018). Kegiatan penambangan data dapat dibedakan bergantung pada hasil yang ingin diperoleh yaitu sebagai berikut:

1. Prediction

Prediction atau prediksi yaitu untuk menemukan pola tertentu dari suatu data yang dapat di ketahui dari variable yang ada pada data yang digunakan untuk memprediksi variable lain yang belum di ketahui nilainya.

2. Description

Description atau deksripsi memahami lebih jauh data yang di amati untuk mengetahui karakteristik dari data tersebut.

3. Klasifikasi

Klasifikasi atau classification yaitu pemrosesan data untuk menemukan fungsi atau model tertentu dari konsep suatu data yang digunakan untuk memisahkan tiap data menjadi kelompok-kelompok tertentu.

4. Asosiasi

Asosiasi berfungsi untuk menemukan kombinasi atau aturan asosiatif dari suatu data yang akan di proses untuk menemukan informasi tentang hubungan variabel satu dengan lainnya.

5. Clustering

Clustering lebih mengarah pada pengelompokan data kedalam kelas yang memiliki kesamaan.

2.2.6.2 Algoritma K-Means

K-Means adalah perhitungan yang digunakan dalam kumpulan diskrit yang mengisolasi informasi ke dalam berbagai kumpulan (Sadewo et al., 2017). Ide dasar algoritma k-means sangatlah sederhana yaitu meminimalkan *Sum of Squared Error* (SSE) antara objek informasi dengan berbagai k *centroid*.

Langkah-langkah melakukan clustering dengan menggunakan metode k-means adalah sebagai berikut:

1. Pilih jumlah cluster K.
2. Inisiasi k pusat cluster ini bisa dilakukan dengan berbagai cara, namun yang paling sering dilakukan adalah dengan cara random.
3. Alokasikan semua data/objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu.

$$D(i,j) = \sqrt{(X_{i1} - X_{j1})^2 + (X_{i2} - X_{j2})^2 + \dots + (X_{in} - X_{jn})^2} \quad \text{----- (2.1)}$$

Sumber : (Sari, 2021)

Rumus 2. 1 Euclidian Distance

Dimana:

$D(I,j)$ = jarak data ke I ke pusat cluster j

X_{ki} = data ke I atribut data ke k

X_{kj} = titik pusat ke j pada atribut ke k.

4. Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang, pusat cluster adalah rata-rata dari semua data/objek dalam cluster tertentu.
5. Hitung kembali pusat cluster yang baru. Jika data tidak ada yang berubah dengan iterasi sebelumnya maka proses clustering dapat dihentikan. Atau jika masih ada data yang berubah maka ulangi kembali langkah ke tiga.

2.2.6.3 Clustering

Clustering adalah suatu metode untuk mencari dan mengelompokkan data yang memiliki kesamaan karakteristik (*similarity*) antara data yang satu dengan data yang lain (Sulastri & Gufroni, 2017). Data-data yang mempunyai kesamaan kriteria akan dikumpulkan kedalam kelompok atau *cluster* yang dibedakan. Partisi data dilakukan dengan mencari nilai jarak terdekat antara data dengan nilai centroid yang telah ditetapkan baik secara random maupun dengan *Initial Set of Centroids*, untuk menentukan nilai centroid berdasarkan K objek yang berurutan (Kusanti & Sutanto, 2021). Alasan dasar dari teknik clustering yaitu mengelompokkan berbagai data atau objek kedalam *cluster* (group) hingga dalam masing-masing *cluster* akan berisikan data yang memiliki kesamaan.

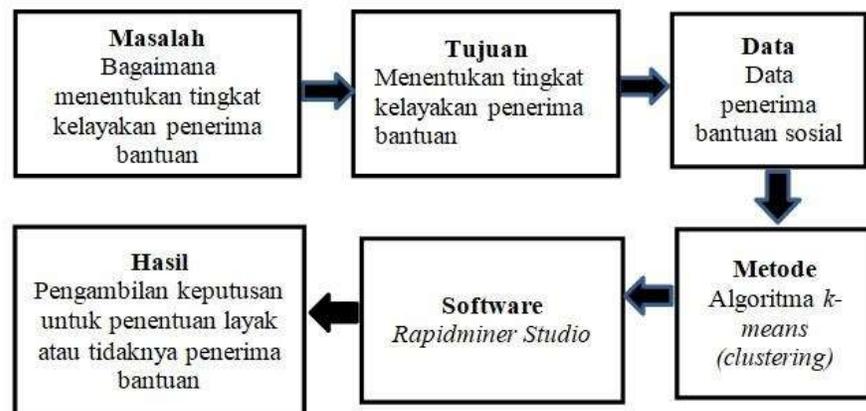
Menurut(Chandrasekar & Kumaran, 2020) *data mining is a technique for gathering information into specific instances or a strategy for sorting mountains of data into significant piles.*

2.2.6.4 Rapidminer

Menurut Satria Wahono dalam jurnalnya (Rony, 2016) Rapidminer merupakan aplikasi open source berlisensi AGPL (GNU Affero General Public License) yang dikembangkan oleh Ralf Klinkenberg, Ingo Mierswa dan Simon Fischer di *artificial intelligence*.

2.3 Kerangka Pemikiran

Berdasarkan permasalahan yang akan diteliti maka kerangka pemikiran dalam penelitian ini adalah:



Sumber: Penulis

Gambar 2. 2 Kerangka pemikiran

Kerangka pemikiran yang terbentuk diawali dengan mengidentifikasi permasalahan yang terjadi lalu memahami tujuan dari penelitian yang akan dilakukan. Data yang digunakan merupakan data penerima bantuan yang telah diperoleh dan memprosesnya dengan menggunakan metode K-Means Clustering

lalu diuji dengan software Rapid Miner Studio sehingga akan didapatkan hasil yang akan dijadikan dasar pengambilan keputusan apakah penerima bantuan dapat dikategorikan layak atau tidaknya menerima bantuan tersebut.

2.4 Penelitian Terdahulu

Penelitian terdahulu bertujuan sebagai referensi dalam sebuah penelitian, maka penelitian terdahulu yang berhubungan dengan penelitian yang dilakukan adalah sebagai berikut:

No	Nama Peneliti, Tahun dan ISSN	Judul Penelitian	Hasil
1	Suhartini ,Ria Yuliani(2021) No. ISSN 2614-8773	Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma KMeans di Dusun Bagik Endep Sukamulia Timur	Penelitian ini menggunakan data mining dengan metode cluster untuk mencari jumlah penduduk miskin Dusun Bagik Endep Sukamulia Timur, Hasil penelitian berupa pengelompokan data penduduk miskin berdasarkan kriteria yang telah ditetapkan.
2	Zainul Aras Z, Sarjono (2016) No. ISSN 2540-8011	Analisis Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan Bedah Rumah Menggunakan Metode Clustering K-Means	Penelitian ini Menggunakan kriteria status kesejahteraan, status pengusaan abngunan tempat tinggal, jenis atap, jenis lantai, jenis dinding dan jumlah indovidu dalam sebuah keluarga untuk menentukan elompok prioritas penerima bantuan bedah rumah bahar utara.

3	Nur Heri Cahyana, Agus Sasmito Aribowo (2018) No. ISSN 9-772301- 936005	Metode Data Mining K-Means Untuk Klasterisasi Data Penanganan Dan Pelayanan Kesehatan Masyarakat	Hasil dari penelitian ini adalah k-Means digunakan untuk pendataan wilayah atau daerah ke dalam beberapa klaster. Untuk mengetahui kondisi kesehatan daerah tersebut.
4	M. Nanda Variestha Waworuntu, Muhammad Faisal Amin(2018) No. ISSN 2406-7857	Penerapan Metode K-Means Untuk Pemetaan Calon Penerima Jamkesda	Penelitian ini menggunakan data mining untuk pemetaan calon penerima jamkesda yang membagi data kedalam 2 klaster. Sebanyak 440 data digunakan sebagai sampel data. Hasil proses data mining adalah didapati 2 cluster yang memenuhi untuk menjadi penerima bantuan yang tepat sasaran.
5	Jani Kusanti, Djoko Sutanto(2021) No. ISSN 2655-9102	<i>Combination of Decision Tree and K-Means Clustering Methods for Decision Making of BLT Recipients in the Covid-19 Period</i>	Penelitian ini menggunakan data mining untuk Pengambilan Keputusan Penerima BLT Masa Covid-19 dengan kriteria Pendapatan (Total Pendapatan/bulan), Pengeluaranpen), Kepemilikan Aset ,dan Tempat Tinggal Hasil proses data mining kriteria komunitas yang berhak menerima data dengan tingkat kesalahan -2,48% < dari toleransi

6	Said Abdul Azis1, Sarjon Defit , Yuhandri Yunus(2021) No. ISSN 2714-8491	Klasterisasi Dana Bantuan Pada Program Keluarga Harapan (PKH) Menggunakan Metode K-Means	Penelitian ini Mengelompokkan data bantuan sosial dengan syarat yang terlah ditetapkan guna pengambilan keputusan pemberian bantuan. Hasil proses data mining menghasilkan kelompok yang terbagi berdasarkan tiga klaster yang ditetapkan.
7	Yunita Ratna Sari, Arby Sudewa, Diah Ayu Lestari, Tri Ika Jaya(2020) No. ISSN 2502-714x	Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer	Penelitian ini menggunakan data mining untuk mengelompokkan jumlah data penduduk miskin dengan kriteria jumlah penduduk miskin, rata-rata lama pendidikan sekolah, dan pengeluaran per kapita yang d disesuaikan (ribu rupiah/tahun) yang dijadikan tolak ukur pengambilan keputusan.

Tabel 2. 1 Penelitian Terdahulu