

BAB II TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

No	Nama Penulis (Tahun)	Isi dan Tujuan	Persamaan	Perbedaan
1.	(Munandar & Sutopo, 2020)	Penelitian ini dilatarbelakangi oleh kesulitan yang dialami koperasi dalam menentukan kelayakan nasabah dalam memberikan kepercayaan melakukan pinjaman atau tidak. Sistem ini akan membantu menentukan nasabah yang layak atau tidak diberikan pinjaman. Beberapa kriteria yang digunakan untuk penentuan diantaranya penghasilan nasabah, besar pinjaman yang diajukan dan lama masa kredit.	Penelitian ini sama dengan penelitian penulis dalam penggunaan algoritma C4.5 berbasis web dalam menentukan kualifikasi kelayakan	Perbedaannya terdapat pada objek penelitian dimana penulis menggunakannya untuk menentukan kelayakan calon penerima bantuan terdampak <i>Covid-19</i> .
2.	(Irmayansyah & Firdaus, 2018)	Tidak akuratnya pelaksanaan program pemberian bantuan kepada penduduk tidak mampu dikarenakan data	Kesamaan yang dimiliki dengan penelitian penulis adalah penggunaan	Perbedaannya yaitu pada sistem informasi yang akan dibuat. Penulis

		yang diterima tidak sesuai dengan keadaan sebenarnya. Penulis menggunakan metode algoritma C4.5 untuk mengklasifikasi menggunakan data berdasarkan sensus ekonomi penduduk yang akan di implementasikan pada sistem informasi berbasis android <i>webview</i> .	algoritma C4.5 untuk mengklasifikasi kelayakan penerima bantuan berdasarkan variabel yang telah ditentukan.	Irmayansyah dan Firdaus membuat sistem informasi berbasis <i>android webview</i> sedangkan yang akan dbuat oleh penulis adalah sistem informasi berbasis <i>web</i> .
3.	(Lestari & Targiono , 2017)	Penelitian ini mengembangkan aplikasi sistem pendukung keputusan penentuan keluarga miskin yang dibuat dengan metode Simple Additive Weighting (SAW) untuk mendukung berbagai macam program pemerintah terkait yang sarannya belum sesuai dengan apa yang diharapkan.	Penulis akan membuat aplikasi yang sama dengan peneliti yaitu sistem pendukung keputusan untuk menentukan kelayakan calon penerima bantuan terdampak covid-19.	Perbedaannya ada pada metode yang digunakan, jika penelitian tersebut menggunakan metode Simple Additive Weighting (SAW) maka penulis menggunakan metode algoritma C4.5 untuk menentukan kelayakan penerima bantuan.

2.2 Kajian Pustaka

2.2.1 Bantuan Terdampak Covid-19

Pemerintah Indonesia telah mengadakan beberapa program guna untuk menguatkan ketahanan sosial serta dorongan ekonomi diantaranya: Program Keluarga Harapan (PKH), Kartu Prakerja, Bansos, Stimulus UMKM Mikro dan Ultra Mikro, serta penurunan tarif listrik serta menaikkan jumlah penerima bantuan akibat dari pandemi *Covid-19* ini (Pramanik, 2020).

2.2.2 Data Mining

Sebuah cara yang berguna untuk mencari pola bermanfaat yang diingkan dan diambil dari sebuah database yang besar atau biasa disebut dengan data mining (Astuti, 2013). Dari seluruh rangkaian proses yang ada pada KDD dalam menemukan pola bermanfaat, data mining menjadi salah satu proses yang sangat penting. KDD atau *Knowledge Discovery in Database* merupakan sekumpulan proses yang digunakan untuk menemukan pengetahuan yang meliputi :

- 1) *Data Cleaning* : merupakan operasi data yang digunakan untuk penghapusan noise, membuang duplikasi data dan juga untuk memeriksa data yang tidak berguna.
- 2) *Data Integration* : hybrid dari macam – macam *database* yang disatukan disatu *database* .
- 3) *Data Selection* : merupakan penyeleksian sekumpulan data set yang perlu dilakukan karena dalam proses analisis tidak semua data atau variable itu diperlukan.
- 4) *Data Transformation* : proses transformasi yang terjadi dan digunakan untuk menghilangkan data yang tidak dipakai hingga sesuai untuk dipergunakan dalam data mining.
- 5) *Data Mining* : sebuah proses pencarian pola terbaik dalam data yang dipilih menggunakan teknik dan metode tertentu seperti klasifikasi, regresi dan lain-lain.

6) *Knowledge Presentation* : merupakan proses penyajian pengetahuan yang telah diperoleh *user* terkait metode yang digunakannya.

Data mining memiliki beberapa metode yang diklasifikasikan berdasarkan tugas yang dilakukan yaitu sebagai berikut :

- *Classification* yaitu cara yang dapat digunakan untuk mengklasifikasikan item data dalam satu kelas dari banyak kelas yang ada.
- *Clustering* yaitu alat untuk menganalisa data untuk menemukan kelompok komponennya secara natural yang didasarkan pada beberapa kesamaan.
- *Association* yaitu penemuan atribut yang seringkali muncul dalam satu waktu.
- *Sequencing* yaitu mengidentifikasi hubungan yang berbeda dalam suatu periode waktu seperti pelanggan yang mengunjungi satu tempat secara berulang kali.
- *Forecasting* yaitu proses perkiraan sebuah nilai dimasa depan menggunakan nilai dari variable lainnya seperti prediksi terjadinya suatu hal dimasa mendatang.

2.2.3 PHP

PHP merupakan bahasa pemrograman yang terpasang pada HTML dan salah satu bahasa pemrograman yang paling banyak digunakan saat ini. Menurut Agus Bahtiar dalam (Nagara & Nurhayati, 2015), PHP telah didesain sedemikian mudahnya agar bisa disisipkan ke halaman HTML. Sebagian besar sintaksnya mirip dengan bahasa lain seperti bahasa C, Java dan Perl namun PHP memiliki beberapa fungsi yang spesifik dengan tujuan untuk memudahkan para programmer merancang web dinamis dengan cepat.

PHP memiliki banyak keunggulan diantaranya :

- a. Memiliki tingkat akses yang lebih cepat.
- b. Memiliki tingkat *lifeecycle* yang cepat

- c. Memiliki tingkat keamanan yang tinggi.
- d. Mampu berjalan di beberapa server, seperti Apache, Microsoft IIS, PWS, AOSserver, phttpd, fhttpd, dan Xitami.
- e. Mampu berjalan di Linux sebagai *platform* sistem operasi utama bagi PHP.
- f. Mendukung ke beberapa *database* yang sudah ada.
- g. PHP bersifat gratis.

2.2.4 Mysql

MySQL merupakan sebuah implementasi dari RDBMS yang didistribusikan secara gratis dengan batasan perangkat lunak hingga tidak boleh dijadikan produk turunan bersifat komersial. Sebagai pelayan data, MySQL mendukung beberapa operasi basis data mulai dari transaksional maupun non-transaksional, oleh karena itu MySQL dikatakan unggul dalam hal unjuk kerja jika dibandingkan dengan yang lainnya. Namun begitu, MySQL juga memiliki kekurangan yaitu tidak adanya jaminan atas reliabilitasterhadap data yang telah disimpan (Polakitan, 2017).

2.2.5 Decision Tree

Decision tree adalah struktur *flowchart* yang memiliki *tree* (pohon), dimana setiap simpul *internal* menandakan suatu tes atribut, setiap cabang mempresentasikan hasil tes, dan simpul daun mempresentasikan kelas atau atribut kelas. *Decision tree* dapat ditelusuri dari simpul akar ke simpul daun yang memegang prediksi kelas. *Decision tree* merupakan suatu metode yang digunakan untuk pengklasifikasian dan prediksi karena memiliki kemudahan dalam interpretasi hasil.

Untuk membentuk *decision tree*, banyak algoritma yang dapat dipakai. Antara lain *ID3*, *CART*, dan *C4.5*. *Decision tree* digunakan untuk mendapatkan informasi yang dapat digunakan untuk mengambil keputusan. (Rismayanti, 2016).

2.2.6 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang banyak dikenal dan digunakan untuk mengklasifikasikan data yang memiliki atribut – atribut numerik dan kategorial. Proses klasifikasi menghasilkan aturan – aturan yang dapat digunakan untuk memprediksi nilai atribut bertipe diskret dari record yang baru. Algoritma C5.4 merupakan pengembangan dari algoritma ID3, dimana pengembangan dilakukan dapat mengatasi *missing* data dan dapat mengatasi data kontinu dan *pruning*.

Algoritma C4.5 untuk membangun pohon keputusan secara umum adalah sebagai berikut :

1. Pilih atribut sebagai akar
2. Buat cabang untuk tiap – tiap nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk setiap cabang hingga semua cabang memiliki kelas yang sama

Untuk memilih atribut cabang, didasarkan pada nilai gain tertinggi dari atribut – atribut yang ada. Untuk menghitung gain, menggunakan rumus sebagai berikut (Elisa, E. 2017) :

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_i Z_i = \sum_i \text{Entropy}(S_i)$$

Dimana :

S : himpunan kasus

A : atribut

N : jumlah partisi atribut A

$|S_i|$: jumlah kasus pada partisi ke- i

$|S^r|$: jumlah kasus dalam S

2.2.7 WEKA (*Waikato Environment for Knowledge Analysis*)

Weka adalah paket *tool machine learning* yang digunakan untuk penelitian, pendidikan dan berbagai aplikasi. WEKA mampu menyelesaikan masalah – masalah data mining di dunia nyata. Perangkat ini ditulis dalam hirarki *class java* dengan metode berorientasi objek dan

dapat berjalan hampir di semua *platform*. Pada aplikasi WEKA tersedia implementasi algoritma – algoritma pembelajaran *state of the art* yang dapat diterapkan pada dataset dari *command line*. WEKA menyediakan *tools* untuk *pre – processing data*, klasifikasi, *regresi*, *clustering*, aturan asosiasi, dan visualisasi.

Ada banyak metode data mining yang dapat dikerjakan atau diselesaikan dengan menggunakan WEKA, salah satunya adalah pohon keputusan (*decision tree*). Pohon keputusan merupakan salah satu metode klasifikasi yang sangat menarik yang melibatkan konstruksi pohon keputusan yang terdiri dari *node* keputusan yang dihubungkan dengan cabang . cabang dari simpul akar sampai ke *node* daun (akhir). Pada *node* keputusan, atribut akan diuji dan setiap hasil akan menghasilkan cabang dan setiap cabang akan diarahkan ke *node* lain ke *node* akhir untuk menghasilkan keputusan. (Pujiono, S., Amborowati, A., & Suyanto, M. 2013).

2.2.8 Web (Website)

Sekumpulan dokumen yang disimpan di internet dan dapat diakses menggunakan HTTP atau *Hyper Text Transfer Protocol* adalah *Website* dimana dapat memberikan informasi berupa text, video, audio, gambar serta animasi atau kombinasi semuanya (Supriyanto, 2007). Pada awalnya *website* merupakan layanan yang menyajikan informasi menggunakan konsep hyperlink dimana memudahkan pengguna internet melakukan penelusuran informasi. Menurut Yuhefizar (1998) *website* adalah suatu metode untuk menampilkan informasi di internet.

Website dibedakan menjadi dua, berdasarkan sifatnya:

1.) *Website Statis*

Website merupakan *website* yang kontennya tidak dapat dirubah oleh pengguna dari browser secara langsung. Pada *website statis* tidak ada database yang ditampilkan, data yang muncul pun tidak berubah-ubah kecuali ada perubahan pada sintaksnya.

2.) *Website Dinamis*

Dalam *website dinamis*, interaksi yang terjadi antara pengguna dan *server* sangat kompleks. Pengguna *website* dapat mengubah konten dari halaman tertentu dengan menggunakan *browser*. Permintaan dari pengguna dapat diproses oleh *server* yang kemudian ditampilkan dalam isi yang berbeda – beda menurut alur programnya. Halaman – halaman *web* tersebut memiliki *database*. *Website dinamis* memiliki data dan informasi yang berbeda – beda tergantung input yang disampaikan *client*. Dokumen yang sampai di *client* akan berbeda dengan dokumen yang ada di *web server*.

Perbedaan *web statis* dan *web dinamis*

- Interaksi antar pengunjung dan pemilik *web*
- Bahasa *script* yang digunakan
- Penggunaan *database*
- Konten

2.3 Pengujian Sistem

Pengujian sistem dilakukan dengan metode *black box*. Metode pengujian *blackbox* mengeksekusi data uji dan memeriksa fungsi sebuah aplikasi. Fokus pengujian terletak pada fungsi sistem (Rizki 2015, dalam Hanifah et al. 2016).

Metode ini digunakan untuk mengetahui apakah perangkat lunak yang dihasilkan berfungsi dengan benar. Pengujian *black box* berusaha menemukan kesalahan dalam kategori berikut:

1. Fungsi – fungsi yang tidak benar ataupun hilang
2. Kesalahan interface
3. Kesalahan dalam struktur data atau akses database eksternal
4. Kesalahan kinerja
5. Inisialisasi dan kesalahan terminasi (Setiawan, G. W. 2011)

Menurut (Julian 2015) pengujian *black box* terdiri dari 10 tipe pengujian, penjelasan dari setiap tipe adalah sebagai berikut:

1. *Equivalence Partition* yaitu membagi inputan menjadi kelas data yang dapat digunakan untuk pengujian.
2. *Boundary Value Analysis* yaitu dengan menyeleksi kasus uji yang menguji batasan nilai input dan merupakan komplemen dari *Equivalence Partitioning*.
3. *Comparison Testing* merupakan tahap menguji setiap versi dengan data yang sama untuk memastikan semua versi menghasilkan keluaran yang sama.
4. *Sample Testing* yaitu dengan melibatkan beberapa nilai yang terpilih dari sebuah ekivalen.
5. *Robustness Testing* yaitu data input yang dipilih diluar dari spesifikasi yang telah didefinisikan dengan tujuan untuk membuktikan jika tidak ada kesalahan
6. *Behavior Testing* yaitu melakukan pengujian berulang kali untuk memastikan tidak ada kesalahan pada sistem
7. *Performance Testing* merupakan tahap evaluasi kemampuan sistem untuk beroperasi dengan benar.
8. *Requirement Testing* merupakan spesifikasi kebutuhan yang terasosiasi dengan perangkat lunak yang diidentifikasi pada tahap spesifikasi kebutuhan dan desain.
9. *Endurance Testing* melibatkan kasus uji yang diulang-ulang dengan jumlah tertentu.
10. *Cause – Effect Relationship Testing* merupakan pembagian spesifikasi kebutuhan menjadi bagian yang memiliki kemungkinan kerja. (Hanifah et al. 2016)