

BAB I

PENDAHULUAN

1.1. Latar Belakang

Rumah yang layak huni menjadi keinginan setiap keluarga namun masyarakat yang berpenghasilan rendah tidak mampu membangun rumah yang layak huni oleh sebab itu salah satu upaya pemerintah memperbaiki rumah masyarakat yang tidak layak huni yaitu dengan program bantuan sosial rumah tidak layak huni atau rutilahu. Penyaluran bantuan rutilahu tidak tepat sasaran karena selama ini tidak ada metode dalam menentukan penerima bantuan rutilahu. Dampak dari penyaluran bantuan rutilahu yang tidak tepat sasaran berimbas pada pembangunan rumah yang tidak selesai. Program rutilahu merupakan upaya pemerintah Kabupaten Bekasi dalam pemerataan pembangunan disetiap Desa dan sasaran dari program rutilahu ditunjukkan kepada masyarakat berpenghasilan rendah atau masyarakat swadaya yang mempunyai pekerjaan dengan penghasilan terbatas, hal ini sesuai dengan peraturan bupati nomer 3 tahun 2019 tentang bantuan rumah tidak layak huni untuk masyarakat berpenghasilan rendah di Kabupaten Bekasi.

Penelitian yang telah dilakukan mengenai klasifikasi penerima bantuan rutilahu dengan menggunakan algoritma C4.5 dan menggunakan teknik *statistic* kemudain diekstraksi data tersebut sehingga menghasilkan klasifikasi data dalam menentukan layak dan tidak layak menerima bantuan bedah rumah (Gunawan, Damanik, parlina & Saputra 2021). Metode yang digunakan dalam penelitiannya menggunakan metode vikor dengan multi kriteria didalam sistem yang kompleks untuk mendapatkan hasil ranking sebagai solusi alternatif (Tumanggor, Haloho, Ramdhani & Nasution, 2018). Metode vikor sebagai metode yang *Multi Criteria Decision Making (MCDM)* untuk menyeleksi lebih dari satu kriteria dan sampel berdasarkan ranking dengan melihat nilai regrets (R) sehingga hasil indeks vikor terkecil merupakan ranking pertama (Manurung, Damanik & Saragih 2021). Peneliti menerapkan metode SAW pada data calon penerima bantuan rutilahu, kemudia data dimasukan ke dalam bahasa pemrograman PHP dan disimpan dalam *database MYSQL* dan nilai prefensi dianggap sebagai nilai akhir (Manurung,

Damanik & Saragih 2021). Peneliti menggunakan dua metode yaitu metode smarter dan metode tophis, metode smarter digunakan sebagai bobot kriteria penting metode tophis digunakan untuk menghitung kriteria bobot yang penting kemudian nilai yang memiliki perhitungan yang tinggi layak untuk mendapatkan bantuan rutilahu (Saepudin, Gustian & Firmansyah 2019). Terakhir peneliti menggunakan metode tophis dalam penelitiannya dengan menentukan kriteria untuk seleksi penerima bantuan rutilahu dan menentukan peringkat kriteria sebagai pilihan kriteria, nilai dari pilihan terbesar merupakan pilihan terbaik (Nalatissifa & Ramdhani 2020).

Pemerintah membuat program rutilahu untuk membantu masyarakat dalam membangun rumah yang layak huni. Dengan adanya rumah yang layak huni masyarakat bisa nyaman dan tenang saat melaksanakan kegiatan sehari-hari. Tidak hanya partisipasi dari pemerintah, Partisipasi masyarakat juga penting dalam pembangunan program rutilahu karena masyarakat itu sendiri akan terlibat baik secara langsung ataupun tidak langsung dalam suksesnya program rutilahu. Dalam hal ini pemerintah Labansari harus mengambil keputusan keluarga mana yang layak untuk mendapatkan bantuan rutilahu dengan mempertimbangkan kriteria yang cocok sesuai dengan peraturan pemerintah. Pemberian program rutilahu kepada masyarakat terkadang belum tepat sasaran karena klasifikasi yang dijalankan masih manual dalam pengolahan data masyarakat sehingga dukungan dari pemerintah belum optimal karena ada kesalahan dalam mengolah data dari calon penerima bantuan rutilahu hal ini dijabarkan oleh Kepala Seksi Kesejahteraan Desa Labansari.

Dari latar belakang diatas penulis mengusulkan penelitian dengan judul penerapan algoritma C4.5 untuk klasifikasi penerima bantuan rutilahu Desa Labansari. Penelitian ini diusulkan karena tidak akuratnya penyaluran bantuan rumah tidak layak huni karena tidak ada metode klasifikasi maka untuk memperkecil kesalahan dalam mengambil keputusan data diklasifikasi menggunakan algoritma C4.5. Dari 15 penerima bantuan rutilahu yang belum tepat sasaran sebesar 33,33%.

1.1. Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan, adapun perumusan masalah adalah sebagai berikut:

1. Bagaimana mengklasifikasi penerima bantuan rutilahu menggunakan algoritma C4.5.
2. Bagaimana nilai akurasi, presisi dan *recall* dalam mengklasifikasi penerima bantuan rutilahu menggunakan algoritma C4.5.

1.2. Tujuan Penelitian

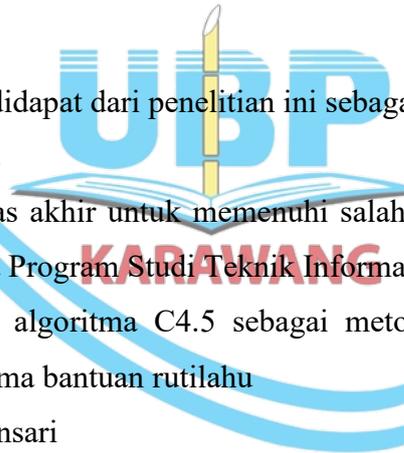
Tujuan dari tugas akhir ini yaitu :

1. Menerapkan algoritma C4.5 untuk mengklasifikasi penerima bantuan rutilahu
2. Mengetahui nilai akurasi, presisi dan *recall* dari data yang diolah menggunakan algoritma C4.5. ★

1.3. Manfaat

Manfaat yang bisa didapat dari penelitian ini sebagai berikut:

1. Bagi mahasiswa
 - a. Sebagai tugas akhir untuk memenuhi salah satu syarat memperoleh gelar sarjana Program Studi Teknik Informatika
 - b. Menerapkan algoritma C4.5 sebagai metode mengklasifikasi data calon penerima bantuan rutilahu
2. Bagi Desa Labansari
 - a. semoga bisa menjadi bahan evaluasi bagi Desa Labansari dalam memilih keluarga yang akan diberikan bantuan rutilahu.



BAB II

TINJAUAN PUSTAKA

2.1. Rumah

Rumah merupakan tempat tinggal yang memiliki banyak fungsi sebagai tempat yang layak untuk manusia (Tias Puspita Ningrum, 2018). Rumah sebagai kebutuhan primer manusia sehingga perlu diperhatikan menyangkut kesejahteraan manusia (Fauzi, 2018). Dari penjelasan diatas dapat disimpulkan bahwa rumah merupakan hunian atau tempat tinggal yang berfungsi untuk pembinaan keluarga. Dalam makna yang lebih luas rumah bukan hanya bangunan tetapi tempat yang memiliki syarat-syarat kehidupan yang layak. Dalam Undang-Undang Dasar 1945 pasal 28H Ayat 1 yang berbunyi memiliki rumah atau hunian yang layak merupakan hak bagi setiap warga negara Indonesia (Ardiansyah et al., 2020). Namun kenyataannya masih banyak rumah yang tidak layak huni karena disebabkan ekonomi masyarakat yang belum merata dengan daerah lain.



Gambar 2. 1 Rumah tidak layak huni
(Sumber: Rumah bapak Apet Bin Taim)

Rumah yang tidak layak huni umumnya berada dipemukiman yang kumuh dan miskin. Hal ini berdampak negatif bagi penghuni rumah karena kurangnya pencahayaan dan penghawaan didalam rumah sehingga mengganggu kesehatan dan produktifitas penghuninya. Bantuan rutilahu bertujuan untuk peningkatan kualitas rumah dengan melihat dari sisi keselamatan bangunan, kesehatan hunian, kecukupan luas bangunan dan persyaratan administratif. Dengan begitu penghuni rumah akan menjadi lebih sehat dan produktif serta sebagian masalah kemiskinan bisa teratasi dengan dibangunnya rumah yang layak huni karena akan timbul rasa nyaman dirumah, peran dan fungsi keluarga akan meningkat dalam memberikan perlindungan, pembimbingan dan pendidikan tidak hanya itu saja kualitas hidup dipemukiman akan meningkat karena rumah yang nyaman. Berdasarkan peraturan bupati nomer 3 tahun 2019

tentang bantuan rumah tidak layak huni untuk masyarakat berpenghasilan rendah di Kabupaten Bekasi yang memiliki syarat dan ketentuan sebagai berikut:

1. Hanya berlaku di Kabupaten Bekasi
2. Penduduk ber-KTP Bekasi
3. Memiliki tanah dengan hak yang sah
4. Memiliki dan menempati satu-satunya rumah
5. Kondisi pada bagian atap, lantai dan dinding tidak layak.
6. Pembangunan rumah baru pengganti rumah rusak
7. Pembangunan rumah diatas kavling tanah matang (KTM)
8. Ukuran tanah paling luas 45 meter persegi
9. Terdampak bencana skala Kabupaten
10. Terdampak program pemerintah Kabupaten
11. Rumah tunggal atau
12. Rumah deret

Bentuk bantuan rutilahu berupa uang tunai yang diberikan kepada Lembaga Pemberdayaan Masyarakat (LPM) atau Badan Keswadayaan Masyarakat (BKM) yang digunakan untuk membeli bahan bangunan dan upah kerja. Bantuan rutilahu diawasi oleh Pejabat Pembuat Komitmen (PPK) yang mengawasi semua kegiatan rutilahu mulai dari pengusulan, penetapan lokasi, penyiapan masyarakat, penetapan bantuan rutilahu, pencairan, penyaluran, pemanfaatan dan pelaporan.

2.1.1. Prosedur Pengusulan Rumah

Prosedur pengusulan bantuan rutilahu diatur dalam peraturan bupati nomer 3 tahun 2019 tentang bantuan rumah tidak layak huni untuk masyarakat berpenghasilan rendah di Kabupaten Bekasi sebagai berikut :

1. Masyarakat yang memenuhi persyaratan bantuan rutilahu diajukan oleh rt atau rw kepada LPM sebagai daftar calon penerima bantuan kemudian ditetapkan oleh bupati Bekasi sebagai penerima bantuan rutilahu.
2. Selanjutnya LPM membuat proposal yang berisi jumlah rumah tidak layak huni dan jumlah bantuan rutilahu.
3. Dinas Perumahan Rakyat, kawasan permukiman dan pertahanan, tim teknis melakukan pemeriksaan data kemudian ditetapkan oleh bupati Bekasi sebagai daftar penerima bantuan rutilahu

2.1.2. Prinsip Kegiatan Pembangunan Rumah

Prinsip pembangunan rutilahu diatur dalam peraturan bupati nomer 3 tahun 2019 tentang bantuan rumah tidak layak huni untuk masyarakat berpenghasilan rendah di Kabupaten Bekasi sebagai berikut :

1. Keselamatan bangunan seperti standar kualitas setiap komponen atap, lantai dan dinding bangunan
2. Kesehatan penghunian memenuhi standar kecukupan cahaya, sirkulasi udara dan sarana MCK
3. Kecukupan luas bangunan memenuhi standar ruang gerak manusia untuk kenyamanan bangunan
4. Persyaratan administrasi seperti dibangun pada lokasi yang ditentukan, dibangun pada tanah sendiri bukan tanah sengketa dan rekomendasi Kepala Desa

2.2. Algoritma C4.5

Algoritma C4.5 merupakan pohon yang memiliki struktur dimana terdapat simpul untuk mendeskripsikan atribut yang akan diuji dan kelas digambarkan dengan daun (Widayu et al., 2017). Termasuk algoritma yang bisa dipakai untuk operasi klasifikasi data (Rizal Amegia Saputra, Sri Wasiyanti, 2020). Algoritma C4.5 satu dari banyak metode klasifikasi yang representasi struktur pohon, setiap simpul direpresentasikan sebagai atribut dan setiap daun direpresentasikan sebagai kelas (Anam & Santoso, 2018). Algoritma C4.5 salah satu teknik data mining yang mengolah variabel atribut didalam data sehingga mampu mengklasifikasi data terkait pemberian bantuan sosial rutilahu kepada masyarakat. Menggunakan konsep *gain* dan *entropy* untuk menentukan nilai informasi terbesar sebagai *gain* dan *entropy* yang optimal. Nilai *gain* tertinggi atau memiliki nilai tinggi maka dijadikan simpul pohon keputusan berdasarkan perhitungan *entropy* yang nantinya menjadi poros klasifikasi. Klasifikasi teknik dalam melihat kelakuan dan atribut berdasarkan kelompok yang sudah didefinisikan.

Teknik klasifikasi menghasilkan data baru pada klasifikasi dengan memanipulasi data klasifikasi sehingga menghasilkan sejumlah aturan baru. Algoritma C4.5 diciptakan oleh J. Rose Quinlan yang merupakan pengembangan dari algoritma ID3 (Mardi, 2017). Kinerja dari algoritma C4.5 sangat baik apabila menggunakan *shuffled sampling* dengan *criterion gain ratio*, *shuffled sampling* memiliki kinerja yang bagus ketika diterapkan dengan tipe data *integer* (Mutrofin et al., 2020). Teknik *shuffle sampling* dengan kriteria *gain ratio* terbukti memiliki kinerja terbaik untuk model klasifikasi algoritma C4.5 (Ismanto & Novalia, 2021). Kedua Pernyataan tersebut membuktikan bahwa Algoritma C4.5 memiliki kinerja yang baik dalam mengklasifikasi data. Algoritma C4.5. Alur dari algoritma C4.5 dimulai dari menyiapkan data kemudian mengolah data menggunakan algoritma C4.5 maka akan menghasilkan pohon keputusan dan rule (Rosdiana & Rismayana, 2018). Termasuk metode pohon keputusan yang merepresentasikan dari aturan klasifikasi dalam membentuk pohon keputusan (Anam & Santoso, 2018).

Keunggulan dari algoritma C4.5 memiliki keunggulan dalam data numerik, data kategori dan nilai hilang didalam atribut yang diolah kontinyu (Firstdly, 2018). Selain itu kelebihan dari algoritma C4.5 mampu menginterpretasikan pohon keputusan dengan mudah, akurasi dan efisien untuk menyelesaikan atribut dengan tipe diskret dan numerik (Muhamad et al., 2019). Perhitungan algoritma C4.5 dibagi menjadi dua tahap perhitungan pertama menghitung nilai *entropy* kedua menghitung nilai *gain* (Rizal Amegia Saputra, Sri Wasiyanti, 2020). Algoritma C4.5 membangun pohon keputusan dari atas sampai bawah, akar berada diposisi paling atas dan posisi paling bawah disebut daun, langkah-langkah untuk membuat pohon keputusan secara umum sama (Rosdiana & Rismayana, 2018).

1. Menghitung keseluruhan data berdasarkan atribut tertentu
2. Memilih atribut *node*
3. Membuat cabang *node* dari anggota *node*
4. Menentukan *node* yang terbentuk berdasarkan nilai *entropy*
5. Jika anggota *node* bernilai *entropy* lebih dari nol maka ulangi proses awal

2.2.1. Entropy

Entropy merupakan ukuran informasi yang berisikan teori untuk mengetahui *impurity* dan *homogeneity* berdasarkan karakteristik dari data yang terkumpul (Arifin & Fitriana, 2018). Perhitungan pertama *entropy* yaitu menghitung semua total nilai yang ada dan dilanjutkan dengan nilai disetiap atribut (Muhammad et al., 2019). Perhitungan *entropy* bertujuan untuk mengetahui seberapa informatif dari atribut dalam menghasilkan atribut baru (Muhammad et al., 2019). Perhitungan *entropy* merupakan perhitungan pertama sebelum perhitungan lainnya karena nilai *entropy* akan menjadi dasar pada perhitungan selanjutnya (Widayu et al., 2017). Nilai *entropy* pertama yang dihitung adalah nilai keseluruhan data kemudian menghitung semua nilai dari setiap atribut yang ada sampai tidak ada lagi nilai yang tersisa (Suntoro & Indah, 2017). Setiap atribut dihitung untuk diketahui nilai *entropy* pada masing-masing atribut. Perhitungan *entropy* menggunakan rumus seperti dibawah ini.

$$Entropy (S) = \sum_{n=1}^n - p_i * \log^2 p \quad (2.1)$$

Keterangan

S : Himpunan Kasus

n : Jumlah partisi S

p_i : Proporsi dari S_i terhadap S

Secara matematis suatu *entropy* dengan persamaan n adalah keseluruhan atribut target dan P_i dinyatakan jumlah dari sampel kelas $n=1$ (Suntoro & Indah, 2017). Rumus diatas digunakan untuk menghitung nilai *entropy* pada atribut dan menentukan ukuran informatif pada tiap atribut. nilai *entropy* juga digunakan untuk mengukur informasi atribut berdasarkan jumlah informasi. Setelah semua nilai *entropy* didapatkan maka bisa melakukan perhitungan selanjutnya (Arifin & Fitriana, 2018). nilai *entropy* dengan nilai 0 maka akan dijadikan cabang atau nilai yang paling kecil dari semua atribut (Kurniasari & Fatmawati, 2019). *Entropy* menjadi parameter dalam mengukur *heterogen* (keberagaman) sampel data, semakin banyak keberagaman data maka nilai *entropy* akan semakin besar (Suntoro & Indah, 2017)

2.2.2. Informasi Gain

Informasi *gain* merupakan informasi berdasarkan hasil perhitungan nilai *entropy* pada setiap atribut atribut, baik secara observasi maupun disimpulkan berdasarkan set data (Arifin & Fitriana, 2018). Informasi *gain* menggunakan metode seleksi fitur dengan diperoleh nilai akurasi yang baik (Muqorobin et al., 2019). Tujuan menghitung nilai *gain* yaitu untuk mengurangi dimensi *vektor* (Setiyorini & Asmono, 2019). Kelebihan dari Informasi *gain* yaitu

mampu menghilangkan fitur yang tidak relevan dan mampu mengurangi dimensi dari *vektor* sebelum klasifikasi (Candana et al., 2021). Informasi *gain* menghasilkan data baru sehingga atribut tersebut akan lebih optimal dibandingkan dengan data yang belum dilakukan perhitungan informasi *gain* (Muhamad et al., 2019). Nilai dari informasi *gain* sangat berpengaruh terhadap kelasnya (Muqorobin et al., 2019). Pemilihan atribut didasarkan pada nilai *gain* tertinggi dari atribut yang ada (Elisa, 2017). Perhitungan *gain* menggunakan rumus seperti dibawah ini:

$$Gain(S,A) = Entropy(S) \sum_{n=1}^i \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2.2)$$

Keterangan:

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|S_i| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

Setelah semua nilai *entropy* terhitung selanjutnya bisa menghitung nilai *gain* dan nilai *gain* digunakan untuk mengukur efektifitas atribut diproses klasifikasi. Dengan menggunakan kedua rumus tersebut maka hasil nilai tersebut bisa diproses algoritma C4.5 dan diproses pohon keputusan. Pohon keputusan digunakan dengan tujuan pengubahan data menjadi pohon keputusan Penerapan algoritma C4.5 dapat mengklasifikasi calon penerima bantuan sosial rutilahu dengan cara membuat pohon keputusan berdasarkan data yang ada untuk menerima bantuan sosial rutilahu.

2.3. Klasifikasi

Klasifikasi merupakan pengelompokan data atau fitur sesuai dengan karakteristik data dan kelas tersebut (Aji Prasetya Wibawa et al., 2018). Klasifikasi merupakan metode untuk menemukan model sehingga data tersebut bisa dikelompokkan sesuai dengan model dan fungsi (Lhokseumawe et al., 2020). klasifikasi menemukan fungsi dan model yang dapat membedakan dan menjelaskan konsep data tersebut dan memiliki aturan yang matematis (Lhokseumawe et al., 2020). Algoritma C4.5 merupakan algoritma yang sangat baik untuk mengklasifikasi data, algoritma ini dipilih karena paling baik dalam melakukan dan pengujian klasifikasi (Ismanto & Novalia, 2021). Penelitian ini menggunakan algoritma C4.5 untuk mengklasifikasi data tersebut. Prose klasifikasi dibagi menjadi dua, yang pertama yaitu learned model atau membangun model berdasarkan kelas data serangkaian dan yang kedua yaitu supervised learning atau penentuan kelas satu atribut (Prihatmono & Watratan, 2019).

Metode klasifikasi yang dihasilkan oleh algoritma C4.5 berupa struktur pohon atau pohon keputusan yang merepresentasikan node sebagai atribut, cabangnya sebagai nilai atribut dan daun sebagai kelas (Elisa, 2017). Dalam data mining klasifikasi dibuat dari data dan kelas yang telah ditentukan dan dibagi menjadi beberapa kategori yang berbeda (Aji Prasetya Wibawa et al., 2018). Algoritma C4.5 merupakan teknik klasifikasi yang banyak digunakan dan hasilnya didapatkan berdasarkan aturan yang ada (Romli & Zy, 2020). Dengan teknik klasifikasi akan menghasilkan nilai akurasi, presisi dan *recall* berdasarkan data yang sudah ada sehingga pemberian bantuan rutilahu bisa menjadi tepat sasaran. Klasifikasi memproses dan menempatkan masing-masing objek ke dalam suatu set objek kategori (Romli & Zy, 2020). Proses klasifikasi menghasilkan aturan yang memprediksi dari nilai atribut dengan bagian-bagian yang terpisah dari data yang baru (Elisa, 2017). atribut yang memenuhi kriteria bisa digunakan untuk klasifikasi sehingga menghasilkan arsitektur optimal dan klasifikasi yang akurat (Setiyorini & Asmono, 2019).

2.4. Python

Python merupakan bahasa pemrograman berorientasi objek dinamis, mudah dimengerti dan bisa untuk berbagai macam keperluan (Ngantung & Pakereng, 2021). Python merupakan bahasa pemrograman tingkat tinggi (*high level programming language*) yang diciptakan oleh Guido Van Rossum pada tahun 1990 di Belanda (Gumilar et al., 2021). Bahasa pemrograman python berbasis *open source* sehingga bisa digunakan secara gratis yang didukung dengan *developer productivity*, *componen integration*, *software quality* dan *program portability*

(Gumilar et al., 2021). Keunggulan yang dimiliki oleh bahasa pemrograman python sebagai berikut:

1. Pengkodean lebih sedikit karena dikembangkan dengan cepat
2. Bisa di semua platform
3. Sistem memori dikelola secara otomatis
4. Bersifat *Object Oriented Programming (OOP)*



Gambar 2. 2 Logo bahasa pemrograman python
(Sumber: python.org)

Bahasa pemrograman python bersifat multiguna yang berfokus keterbacaan setiap kode, python mampu menyatukan kemampuan didetail sintaks kode karena didukung dengan pustaka yang komprehensif (Zuraidah et al., 2021). Alur pengkodean python cenderung lebih sedikit oleh karena itu banyak orang yang menyukai bahasa pemrograman python, penulisannyapun lebih sedikit dan lebih cepat. Fitur yang dimiliki python cukup lengkap (Prihatmono & Watratan, 2019).

1. Memiliki modul dengan pustaka yang luas
2. Tata bahasa mudah dipelajari
3. Pengaturan layout yang mudah sehingga memudahkan pembacaan dan pengecekan kode.
4. Bisa dibangun menggunakan bahasa C/C++

2.5. Weka (Waikato Environment For Knowledge Analysis)

Weka merupakan perangkat lunak dengan menggunakan bahasa java yang dikembangkan di Universitas Waikato, New Zealand (Faid, Jasri & Rahmawati 2019). Weka merupakan aplikasi untuk mengolah data dengan menggunakan metode *data mining* yang bisa dipasang disemua sistem operasi yang mendukung *Java Virtual Machine* (Dasa Putri & Kunci 2019). Dari penjelasan diatas dapat disimpulkan bahwa weka merupakan perangkat lunak yang mampu mengolah data menggunakan metode *data mining*, singkatan dari dari weka yaitu *Waikato Environment For Knowledge Analysis*. Weka berbasis *open source software* sehingga bisa digunakan dan dimodifikasi secara gratis karena berlisensi *General Public License* (GNU) (Dasa Putri & Kunci 2019). Algoritma yang dimiliki oleh perangkat lunak ini cukup banyak seperti *machine learning* untuk *data mining* dan juga memiliki banyak *tools* dalam pengolahan datanya seperti *pre-processing*, *classification*, *association rules* dan *visualization* (Faid et al. 2019).



Gambar 2. 3 Logo aplikasi weka
(Sumber: waikato.github.io)

Aplikasi weka digunakan pada penelitian ini karena memiliki keunggulan dibandingkan dengan aplikasi pengolah data lainnya. Pengolahan data yang dilakukan oleh aplikasi weka dengan cara menguraikan kedalam bentuk excel yang menghasilkan bentuk keputusan (Hasugian 2018). Hasil data yang diolah menggunakan aplikasi weka akan diuji keakuratannya berdasarkan algoritma C4.5 sehingga diperoleh nilai *entropy* dan *gain* dari variabel yang ada (Harahap 2021).

2.5.1. Langkah-Langkah Menggunakan Aplikasi Weka

1. Persiapkan *file* yang akan digunakan dengan *format csv*
2. Buka aplikasi weka
3. Pilih tombol *explorer* maka akan muncul halaman kosong
4. Pilih *open file* untuk mencari file csv kemudian pilih *open*
5. Pilih *classify* dan pilih J48 untuk algoritma C4.5
6. Klik *start* untuk memulai proses

2.6. Google Colaboratory

Google colaboratory atau google colab merupakan penyedia layanan gratis yang dibuat oleh google dengan bentuk *executable document* sehingga bisa digunakan menulis, merubah, menyimpan dan membagikan kode program di google drive (Prasanta & Pranata 2021). Google colab dikeluarkan oleh Google Internal Research yang memberikan berbagai pustaka seperti *numpy*, *opencv*, *keras*, *tensorflow* dan lain-lain yang bisa diakses secara bebas melalui browser (Maulidyah 2020). Kelebihan yang dimiliki google colab (Sholado 2020)

1. Dapat diakses secara gratis hanya perlu koneksi internet.
2. Perincian yang ada di google colab cukup baik yang mendukung GPU Nvidia, kapasitas RAM 13 GB, kapasitas penyimpanan 130 GB dan tidak perlu melakukan konfigurasi.
3. Mudah dibagikan karena mendukung dengan google drive dan github serta bisa membuat kode secara bersamaan bahkan dengan menggunakan gawai sekalipun.



Gambar 2. 4 Logo google colaboratory
(Sumber: colab.research.google.com)

Google colab memberikan kapasitas GPU sebesar 12 GB yang mendukung Nvidia, dengan dukungan tersebut proses data yang dilatih berkali-kali dan proses yang dilakukan lebih cepat diperangkat komputer dengan perincian rendah (Prasanta & Pranata 2021). Selain bisa menggunakan bahasa python, google colab mendukung perintah git, pip dan sed. Git adalah fasilitas *version control system* yang digunakan untuk membangun perangkat lunak secara bersamaan. Pip adalah paket python yang menerjemahkan perintah yang ada di

windows dan python interpreter dan sed adalah *stream editor* yang digunakan untuk merekayasa kode tanpa membuka dari berkas sumber (Maulidyah 2020).

2.6.1. Langkah-Langkah Menggunakan Google Colaboratory

1. Pastikan laptop atau komputer anda tersambung dengan *internet*
2. Buka google drive kemudian pilih baru atau tambah, pilih lainnya, pilih google colaboratory maka akan tampil halaman baru google colaboratory
3. Pilih hubungkan sampai muncul keterangan ram dan disk
4. Ketikkan kode yang dibutuhkan
5. Pilih simpan maka kode yang diketikkan akan tersimpan di google drive

2.7. Confusion Matrix

Confusion matrix merupakan metode dalam menghitung akurasi yang dihasilkan diperhitungan data mining (Normawati & Prayogi, 2021). Pengujian dilakukan untuk mengetahui tingkat akurasi, presisi dan *recall* perhitungan dan memastikan perhitungan yang dilakukan benar dan sesuai (Jeklin, 2017). Akurasi untuk mengetahui jumlah klasifikasi yang benar berdasarkan perhitungan total klasifikasi, presisi untuk mengidentifikasi perkiraan jumlah kasus *positif* yang benar dan *recall* untuk mengetahui kasus *positif* dan diidentifikasi benar. Hasil dari perhitungan dianalisa dan divalidasi secara manual, nilai dari hasil perhitungan sama atau tidak dengan nilai *confusion matrix*. Nilai dari *confusion matrix* terbentuk dari *true positif* dan *true negatif* mampu menggambarkan kinerja klasifikasi dari algoritma C4.5 (Pratiwi et al., 2021). Berapa banyak yang mendapatkan bantuan sosial rutilahu dan *true negatif* didapatkan dari berapa banyak yang tidak mendapatkan bantuan sosial rutilahu. Adapun matriks dan rumus akurasi, presisi dan *recall* ada dibawah ini:

Tabel 2. 1 *Confusion matrix*

Prediksi	Fakta	
	<i>Positif</i>	<i>Negatif</i>
<i>Positif</i>	<i>True positif (TP)</i>	<i>False Positif (FN)</i>
<i>Negatif</i>	<i>False Negatif (FN)</i>	<i>True negatif (TN)</i>

1. Akurasi merupakan perbandingan kasus data yang diklasifikasikan benar dengan jumlah seluruh data kasus. Perhitungan nilai akurasi dengan cara menjumlahkan nilai *True positif (TP)* dengan *True negatif (TN)* dan dibagi dengan jumlah *True positif (TP)* ditambah *False negatif (FN)* ditambah *False negatif (FN)* ditambah *True negatif (TN)* seperti persamaan dibawah ini.

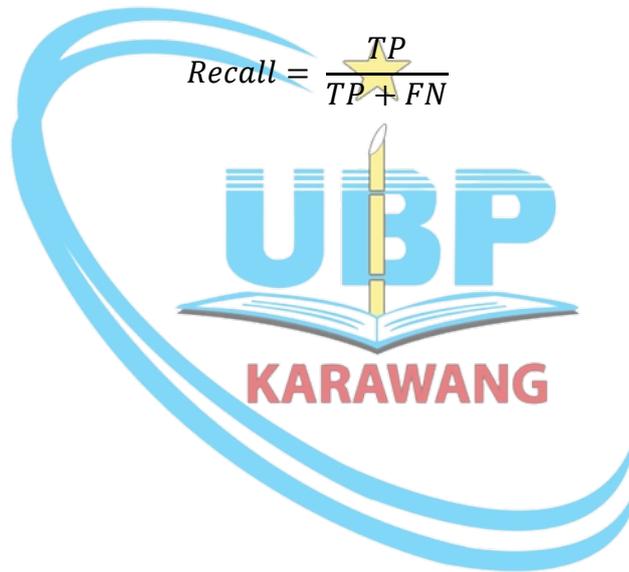
$$Akurasi = \frac{TP + TN}{TP + FN + FP + TN} \quad (2.3)$$

2. Presisi merupakan tingkat kedekatan kasus *positif* yang benar. Nilai presisi didapatkan dengan membagi nilai *True positif (TP)* dengan *True positif (TP)* ditambah *False positif (FP)* seperti persamaan dibawah ini.

$$Presisi = \frac{TP}{TP + FP} \quad (2.4)$$

3. *Recall* merupakan seberapa banyak data *positif* yang teridentifikasi benar. Nilai *recall* dihitung dengan membagi *True positif (TP)* dengan *True positif (TP)* ditambah dengan *True negative (TN)* seperti persamaan dibawah ini.

$$Recall = \frac{TP}{TP + FN} \quad (2.5)$$



2.8. Penelitian Terkait

Penelitian terkait digunakan sebagai dasar dalam menyusun penelitian ini. Penelitian terkait digunakan juga untuk mengetahui hasil dari penelitian sebelumnya.

Tabel 2. 2 Penelitian terkait

No.	Judul Penelitian	Metode	Hasil
1.	Penerapan data mining menggunakan algoritma C4.5 dalam menentukan kelayakan penerimaan bantuan bedah rumah pada desa tiga dolok (Indra Gunawan, Bahrudi Efendi Damanik, Iin Parlina dan Widodo Saputra, 2021)	Menggunakan metode algoritma C4.5	<i>Rules</i> yang terbentuk menjadi 3 aturan klasifikasi layak bedah dan tidak layak bedah. <i>rules</i> akan membuktikan algoritma C4.5 mampu menentukan kelayakan penerima bantuan bedah rumah
2.	Penerapan Metode VIKOR Dalam Penentuan Penerima Dana Bantuan Rumah Tidak Layak Huni (Hotmoko Tumanggor, Mardiana Haloho, Putri Ramadhani, Surya Darma Nasution, 2018)	metode yang digunakan dalam penelitiannya menggunakan metode Vikor dengan multi kriteria	nilai alternatif yang menjadi rangking pertama dengan hasil -0,5
3.	Penerapan Metode VIKOR Dalam Penentuan Penerimaan Dana Bantuan Rumah Tidak Layak Huni (Hotben Manurung, Irfan Sudahri Damanik, Ilham Syahputra Saragih, 2021)	metode vikor sebagai metode yang <i>Multi Criteria Decision Making (MCDM)</i>	Indeks terkecil dari metode vikor merupakan peringkat pertama dengan nilai 0,071428571
4.	Sistem pendukung keputusan dengan <i>Simple Additive weighting</i> dalam pemilihan calon penerima bantuan Rumah Tidak Layak Huni (Sudin	menerapkan metode <i>SAW</i>	Perhitungan metode SAW menunjukkan hasil peringkat pertama dengan nilai 86,67

No.	Judul Penelitian	Metode	Hasil
5.	Saepudin, Dudih Gustian, Heri Firmansyah, 2019) Sistem penunjang keputusan penentuan penerima dana bantuan rumah tidak layak huni dengan metode smarter dan topsis pada Desa Rawakalong (Hendri Ardiansyah, Mochamad Bagoes Satria Junianto, Syaeful Machfud, 2020)	menggunakan dua metode yaitu metode smarter dan metode topsis.	Hasil dari perhitungan dengan nilai tertinggi yaitu 0.8039 yang akan menjadi prioritas utama
6.	Penentuan kelayakan penerima bantuan rutilahu di Desa Labansari menggunakan algoritma C4.5 (Ahmad Zaelani, 2022)	Menggunakan metode algoritma C4.5	

