

NASKAH PUBLIKASI

**SISTEM PENGKLASIFIKASIAN PEMILIHAN PENERIMA
BERAS MISKIN (RASKIN) MENGGUNAKAN METODE *NAÏVE*
*BAYES CLASSIFIER (NBC)***

(Studi Kasus: Kelurahan Tambakmerang, Girimarto, Wonogiri)

Program Studi Teknik Informatika



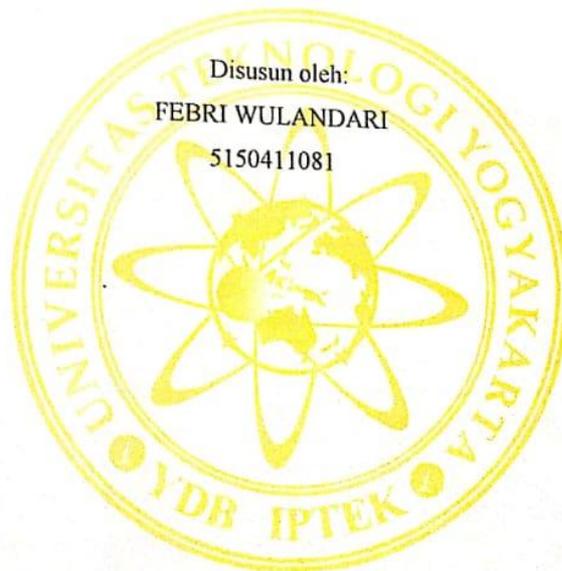
Disusun oleh:

FEBRI WULANDARI

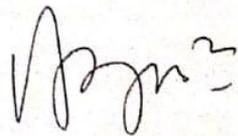
5150411081

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN ELEKTRO
UNIVERSITAS TEKNOLOGI YOGYAKARTA
2019**

NASKAH PUBLIKASI
SISTEM PENGKLASIFIKASIAN PEMILIHAN PENERIMA
BERAS MISKIN (RASKIN) MENGGUNAKAN METODE *NAÏVE*
BAYES CLASSIFIER (NBC)
(Studi Kasus: Kelurahan Tambakmerang, Girimarto, Wonogiri)



Pembimbing



Agus Sujarwadi, S.Kom., M.T.

Tanggal: 23. 8. 19

SISTEM PENGKLASIFIKASIAN PEMILIHAN PENERIMA BERAS MISKIN (RASKIN) MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER (NBC)

(Studi Kasus: Kelurahan Tambakmerang, Girimarto, Wonogiri)

Febri Wulandari, Agus Sujarwadi²

¹Program Studi Teknik Informatika, Fakultas Teknologi Informasi & Elektro

²Program Studi Sistem Informasi, Fakultas Teknologi Informasi & Elektro

Universitas Teknologi Yogyakarta

Jl. Ringroad Utara Jombor Sleman Yogyakarta

Email: febriwulandari51@gmail.com

ABSTRAK

Penyaluran beras miskin atau raskin merupakan kegiatan rutin tiap bulan yang selalu dilakukan pemerintah untuk membantu masyarakat kurang mampu dalam masalah pangan. Awal pelaksanaan raskin bertujuan untuk memperkuat ketahanan pangan rumah tangga terutama rumah tangga miskin. Akan tetapi dalam penentuan kriteria penerima manfaat beras miskin seringkali menjadi persoalan yang rumit. Dalam penentuan kriterianya penerimaannya hanya berdasarkan penilaian subjektifitas saja, sehingga meningkat kecurangan dalam pembagian beras miskin dan mengakibatkan beras tidak rata atau salah sasaran. Oleh karena itu, dilakukan penelitian mengenai penentuan penerima beras miskin. Penelitian ini akan menggunakan metode Naïve Bayes Classifier (NBC). Metode Naïve Bayes Classifier (NBC) merupakan pendekatan yang mengacu pada teorema Bayes yang memprediksi peluang dimasa depan berdasarkan pengalaman di masa sebelumnya. Metode ini merupakan salah satu algoritma klasifikasi yang sederhana namun memiliki akurasi tinggi selain itu metode ini hanya membutuhkan jumlah data pelatihan yang kecil untuk estimasi parameter yang diperlukan dalam proses pengklasifikasian. Data training penerima bantuan beras miskin nanti akan dihitung jumlah labelnya. Dari data tersebut nanti akan dihitung jumlah kriteria yang sama per label, lalu dilakukan perkalian semua variable label. Terakhir adalah membandingkan data hasil per label untuk menentukan kelayakan calon penerima beras miskin (raskin). Berdasarkan hasil penelitian, diharapkan dapat membantu menentukan dalam pengklasifikasian pemilihan penerima beras miskin tanpa adanya kecurangan saat penentuan.

Kata kunci : Beras Miskin (Raskin), Naïve Bayes Classifier (NBC), Calon Penerima.

1.PENDAHULUAN

1.1 Latar Belakang

Kelurahan Tambakmerang adalah salah satu kelurahan yang berada di Desa Tambakmerang, Kecamatan Girimarto, Kabupaten Wonogiri. Kelurahan ini memiliki 9 dusun didalamnya. Kelurahan Tambakmerang dituntut untuk ikut berperan penting dalam mensejahterakan masyarakat salah satunya adalah dalam kelancaran penyaluran beras miskin atau biasa disebut dengan

raskin/ rastra. Raskin merupakan subsidi pangan pokok yang diperuntukkan bagi keluarga miskin sebagai upaya pemerintah dalam menangani krisis pangan. Melalui pelaksanaan program Raskin bersama program bantuan penanggulangan kemiskinan lainnya diharapkan dapat memberikan manfaat yang nyata dalam meningkatkan ketahanan pangan dan kesejahteraan sosial rumah tangga. Penyaluran beras miskin ini terus berjalan sampai dengan saat ini dengan mengikuti kemampuan

subsidi yang dapat diberikan pemerintah kepada keluarga miskin dan perkembangan data keluarga miskin yang ada. Peran pihak kelurahan sangat berpengaruh dalam proses kelancaran penyaluran beras miskin, pihak kelurahanlah yang harus menyalurkan kepada masyarakat penerima manfaat dari beras miskin.

Sasaran penerima manfaat beras miskin adalah masyarakat miskin di dusun – dusun yang terdapat di Kelurahan Tambakmerang yang berhak menerima beras miskin, sebagai hasil seleksi dari pihak kelurahan. Tapi dalam proses penyaluran beras miskin ini memiliki beberapa permasalahan baik dari pihak penyaluran beras miskin atau dari pihak penerima beras miskin. Permasalahan dari pihak penyaluran beras adalah kurangnya kemampuan pihak penyalur dalam mendata penerima beras miskin, mengkategorikan dan mentafsirkan kriteria penerima apakah orang tersebut miskin atau tidak. Kriteria – kriteria yang ditentukan hanya berdasarkan penilaian kira – kira atau subjektif saja. Permasalahan dari pihak penerima manfaat beras miskin adalah rendahnya kesadaran masyarakat akan apa yang seharusnya menjadi haknya atau bukan, banyak masyarakat yang seharusnya tidak berhak mendapatkan bantuan beras miskin tetapi mendapatkan beras miskin, selain itu banyak juga masyarakat yang mendapatkan beras miskin menjual kembali beras tersebut kepada orang lain, menjualnya untuk digantikan oleh bahan pokok lainnya, atau menjualnya karena kualitas rendahnya beras miskin. Hal ini dipengaruhi oleh masyarakat yang mengaku miskin tetapi saat didata ternyata tergolong kepada masyarakat hidup berkecukupan. Salah sasaran inilah yang menjadi permasalahan utama dalam penyaluran beras miskin (raskin) kepada masyarakat miskin yang benar – benar membutuhkan.

Berdasarkan uraian latar belakang permasalahan diatas, untuk mengatasi permasalahan sasaran penerima bantuan beras miskin di Kelurahan Tambakmerang, maka diperlukan suatu sistem pemilihan penerima beras miskin (raskin). Sistem ini dibangun menggunakan metode *Naïve Bayes Classifier (NBC)*. Metode ini adalah salah satu metode algoritma data mining. Sistem ini dipilih karena memiliki algoritma pengklasifikasian yang sederhana dibandingkan dengan algoritma data mining lainnya, selain itu nilai akurasi tinggi. Algoritma ini dapat memberikan prediksi peluang untuk penerima manfaat beras miskin yang sesuai dengan haknya yaitu masyarakat yang tidak mampu.

Dengan menggunakan data sampel dan data pengujian dari Kelurahan Tambakmerang dengan menggunakan kriteria seperti jumlah tanggungan, pekerjaan dan lain sebagainya. Tujuan dari sistem ini nantinya dapat membantu pihak kelurahan dalam mengklasifikasikan pemilihan penerima bantuan beras miskin dan dalam penentuannya menjadi lebih cepat serta lebih akurat.

1.2 Batasan Masalah

Permasalahan yang akan dibahas dan dikaji pada penelitian ini memiliki batasan – batasan yang mencakup:

- a. Kriteria dan atribut untuk sistem pengklasifikasian ini adalah jumlah tanggungan, pekerjaan (wiraswasta, petan transportasi, pedagang, buruh, MRT, dan jasa), tingkat kesejahteraan (miskin, sedang, dan kaya), jenis sarana sanitasi (WC leher angsa, kakus/cemplung, numpang), jenis sumber air (sumur bor/sanyo, ngangsu, kran/hidran umum, PDAM, sumber galian pribadi/sumur), jenis atap (genting, genting lama), jenis dinding (tembok, bambu) dan jenis lantai (keramik, semen, dan lemah).
- b. Output dari sistem berupa klasifikasi layak atau tidak layak menerima beras miskin.

1.3 Tujuan penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

- a. Membangun sebuah sistem klasifikasi pemilihan penerima beras miskin (raskin) sesuai sasaran berdasarkan kriteria – kriteria yang telah ditentukan menggunakan metode *Naïve Bayes Classifier (NBC)*.
- b. Berapa nilai akurasi metode *Naïve Bayes Classifier (NBC)* dalam proses klasifikasi penerimaan beras miskin (raskin).
- c. Berapa perbandingan nilai akurasi metode *Naïve Bayes Classifier (NBC)* menggunakan aplikasi RapidMiner Studio dan sistem yang dibangun.

2. KAJIAN PUSTAKA DAN TEORI

2.1 Landasan Teori

Penelitian terdahulu membahas mengenai penilaian kinerja satpam yang menggunakan sistem “teman menilai teman” yang bisa menimbulkan hasil kinerja yang tidak akurat. Penelitian ini bertujuan untuk menghasilkan sebuah sistem yang mampu melakukan klasifikasi kinerja baik, cukup, dan buruk yang dilihat dari kemampuan, kepribadian dan ketrampilan masing-

masing satpam dengan menggunakan metode Naïve Bayes Classifier. Proses penerapan metode Naive Bayes Classifier untuk klasifikasi kinerja satpam menggunakan perhitungan numerik tiga variabel. Pengujian klasifikasi menggunakan 39 data uji dan menghasilkan prosentase nilai kebenaran sebesar 92,31%, prosentase kinerja baik 20,51%, kinerja cukup 71,79%, kinerja buruk 7,69%. Aplikasi ini bisa digunakan untuk membantu mengetahui kinerja satpam. [11]

Dalam penelitian ini masalah yang dibahas tentang program bersubsidi beras Rastra yang belum optimal dan belum tepat sasaran di karenakan masih banyak keluarga miskin yang tidak mendapatkan beras rastra tersebut. Sehingga diperlukan sebuah metode khusus yang dapat digunakan untuk menanggulangi permasalahan tersebut. Algoritma *Naive Bayes* dapat memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya, pada penelitian ini peneliti mengambil data latih sebanyak 70 data (29 data tidak layak dan 41 data layak) dan sebuah data uji. Variabel yang akan digunakan dalam klasifikasi kelayakan masyarakat penerima beras rastra yaitu Nama Kepala Keluarga, Status PKH dikategorikan menjadi penerima PKH dan tidak menerima PKH, Jumlah Tanggungan yaitu jumlah anak yang masih menjadi tanggungan keluarga, Kepala rumah Tangga dikategorikan menjadi laki – laki dan perempuan, Kondisi Rumah dikategorikan menjadi Batu Permanen, bamboo Anyan, dan Papan, variabel Jumlah Penghasilan dikategorikan menjadi Rendah (<Rp. 1.500.000), Sedang (Rp. 1.500.000 s/d Rp. 2.500.000), dan Tinggi (>Rp. 2.500.000), dan variabel Status Pemilik Rumah dikategorikan menjadi 2 yaitu milik sendiri dan sewa. Hasil penelitian ini menggunakan sebuah data uji menghasilkan nilai Tidak Layak sebesar 0.036322 dan nilai Layak sebesar 0.003567, maka data uji tersebut diklasifikasikan kedalam kelas Tidak Layak. Penerapan data mining dalam menentukan kalsifikasi penerima beras raskin ini dapat digunakan untuk memprediksi layak atau tidak layak sebuah keluarga mendapatkan beras rastra dengan menggunakan algoritma Naive Bayes. [5]

Setiap tahun jumlah pendaftar beasiswa kurang mampu (PPA) bagi mahasiswa yang membutuhkan di Universitas Amikom Yogyakarta terus mengalami peningkatan, namun jumlah yang diterima setiap tahunnya tetap. Oleh karena itu, perlu dikembangkan sistem untuk melakukan data mining dari tumpukan data tersebut yang akan digunakan untuk kepentingan tertentu,

salah satunya adalah untuk menganalisis kelayakan penerima beasiswa agar tepat sasaran. dalam penelitian ini akan dibuktikan kemampuan Naive Bayes Classifier untuk mengklasifikasikan data pendaftar beasiswa yang menginformasikan kelayakan pemberian beasiswa PPA, metode ini dipilih karena mengacu pada teorema Bayes yang mengombinasikan pengetahuan sebelumnya dengan pengetahuan baru. Sehingga merupakan salah satu algoritma klasifikasi yang sederhana namun memiliki akurasi tinggi. Dari hasil pengujian akurasi model dari sistem yang dikembangkan, menghasilkan nilai akurasi terkecil sebesar 64% pada pengujian dengan sampel sebanyak 100 data dan menghasilkan nilai akurasi tertinggi sebesar 97,66% pada pengujian dengan sampel sebanyak 386 data. [1]

Penelitian terdahulu membahas mengenai kelayakan penerimaan beras untuk masyarakat miskin, dalam penelitian ini dibangun sebuah sistem yang mampu membantu untuk membantu menentukan keputusan, sehingga keputusan yang diambil akan lebih objektif. Pada penelitian ini sistem pendukung keputusan digunakan untuk penerimaan beras masyarakat miskin. Algoritma yang digunakan pada sistem pendukung keputusan yaitu algoritma C5.0 dengan model klasifikasi tree. Penerapan algoritma C5.0 pada dataset kelurahan Caringin Wetan dan kelurahan Gunungparang tahun 2015. Model pohon keputusan yang dihasilkan dari penerapan algoritma C5.0 dengan pengolahan data menggunakan SPSS itu dapat digunakan pada aplikasi penentuan penerimaan beras raskin yang akan di gunakan oleh pihak tertentu. Hasil dari penelitian ini adalah dapat menjadi alternatif pemecahan masalah, diantaranya: Sistem pendukung keputusan dengan algoritma C5.0 dibuat agar membantu para pengguna khususnya para petugas kelurahan yang bersangkutan dalam menentukan keputusan mengenai siapa yang benar-benar layak menerima bantuan beras untuk masyarakat miskin. Sistem pendukung keputusan ini dirancang dalam bentuk aplikasi android, sehingga memudahkan para pengguna khususnya para petugas kelurahan yang bersangkutan dalam penggunaannya. Selain itu, karena dibuat dalam aplikasi android, maka informasi yang didapat akan lebih real-time atau bisa didapatkan pada saat itu juga, dan bisa langsung digunakan dimanapun. [6]

Dalam penelitian ini masalah yang dibahas tentang penyamarataan jumlah Raskin yang diterima oleh semua penerima Raskin. Penyaluran Raskin

yang diberikan dalam jumlah yang sama kepada semua penerima Raskin terjadi karena masih dilakukan secara manual sehingga memungkinkan terjadinya penentuan secara subyektif oleh pihak Pemerintah Desa karena belum mengacu pada kriteria yang telah ditentukan. Dalam penelitian ini metode yang digunakan adalah metode SAW (*Simple Additive Weight*). SAW merupakan penjumlahan terbobot, yakni dengan mencari penjumlahan terbobot dari nilai kinerja pada setiap alternatif, pada semua atribut dan membutuhkan proses normalisasi matriks keputusan (X) ke suatu skala yang dapat diperbandingkan dengan semua nilai alternatif yang ada. Hasil dari penelitian ini adalah aplikasi sistem pendukung Keputusan yang dapat membantu kepala desa dalam menentukan jumlah beras yang akan diterima oleh penerima Beras Miskin dan jumlah penerima Beras Miskin didapatkan berdasarkan nilai bobot yang telah ditentukan. Semakin tinggi nilai bobot yang diperoleh oleh penerima maka semakin besar pula jumlah Beras Miskin yang diterima. [3]

2.2 Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstrasi dan mengenali pola yang penting atau menarik dari data yang terdapat pada basis data. Data mining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery Databases* (KDD). [10]

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan machine learning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Menurut Garnet Group, Data Mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan Teknik pengenalan pola seperti teknik static dan matematika. [8]

2.3 Klasifikasi (Classification)

Klasifikasi (*Classification*) proses untuk menyatakan suatu objek ke salah satu kategori yang sudah didefinisikan sebelumnya. Klasifikasi

(Classification) merupakan proses untuk menemukan sekumpulan model yang menjelaskan dan membedakan kelas-kelas data, sehingga model tersebut dapat digunakan untuk memprediksi nilai suatu kelas yang belum diketahui pada sebuah objek. Untuk mendapatkan model, kita harus melakukan analisis terhadap data latih (training set). Sedangkan data uji (test set) digunakan untuk mengetahui tingkat akurasi dari model yang telah dihasilkan. Dalam proses klasifikasi pohon keputusan tradisional, fitur (atribut) dari tupel adalah kategorikal atau numerikal. Biasanya definisi ketepatan nilai (point value) sudah didefinisikan di awal. Pada banyak aplikasi nyata, terkadang muncul suatu nilai yang tidak pasti. Klasifikasi dapat digunakan untuk memprediksi nama atau nilai kelas dari suatu objek data. [4]

2.4 Metode Naïve Bayes Classifier

a. Teorema Bayes

Naïve Bayes Classifier adalah penggolongan menggunakan statistic sederhana berdasarkan teorema bayes yang mengasumsikan bahwa keberadaan atau ketiadaan dari suatu kelas dengan fitur lainnya. [7]

Dalam Bayes (terutama Naïve Bayes), maksud independensi yang kuat pada fitur adalah bahwa sebuah fitur pada sebuah data tidak berkaitan dengan ada atau tidak adanya fitur lain dalam data yang sama. Contohnya, pada kasus klasifikasi hewan dengan fitur penutup kulit, melahirkan, berat dan menyusui. Dalam dunia nyata, hewan yang berkembang biak dengan cara melahirkan dipastikan juga menyusui. Disini ada ketergantungan pada fitur menyusui karena hewan yang menyusui biasanya melahirkan, atau hewan yang bertelur tidak menyusui. Dalam Bayes, hal tersebut tidak dipandang sehingga masing-masing fitur seolah tidak memiliki hubungan apapun. Prediksi Bayes didasarkan pada teorema Bayes dengan formula umum seperti dalam persamaan 2.1. Persamaan 2.1 sebagai berikut: [4]

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} \dots\dots\dots(2.1)$$

Keterangan:

$P(H | E)$ = Probabilitas akhir bersyarat (conditional probability) suatu hipotesis H terjadi jika diberikan bukti (evidence) E terjadi.

$P(E | H)$ = Probabilitas sebuah bukti E terjadi akan mempengaruhi hipotesis H.

- P(H) = Probabilitas awal hipotesis H terjadi tanpa memandang bukti apapun.
- P(E) = Probabilitas awal bukti E terjadi tanpa memandang hipotesis atau bukti yang lain.

Ide dasar dari aturan Bayes adalah bahwa hasil dari hipotesis atau peristiwa (H) dapat diperkirakan pada beberapa bukti (E) yang diamati. Ada beberapa hal penting dari aturan Bayes yaitu:

1. Sebuah probabilitas awal/priori H atau PH adalah probabilitas dari suatu hipotesis sebelum bukti diamati.
2. Sebuah probabilitas akhir H atau (PH) adalah probabilitas dari suatu hipotesis setelah bukti diamati.

Teorema Bayes juga bisa menangani beberapa bukti, misalnya ada E₁, E₂, dan E₃ sehingga akhir untuk hipotesis (H) dapat dihitung dengan cara seperti dalam Persamaan 2.2. Persamaan 2.2 sebagai berikut: [4]

$$P(H|E_1, E_2, E_3) = \frac{P(E_1, E_2, E_3 | H) * P(H)}{P(E_1, E_2, E_3)} \dots \dots \dots (2.2)$$

Karena asumsi yang digunakan untuk bukti adalah independen, bentuk di atas dapat diubah menjadi seperti dalam Persamaan 2.3. Persamaan 2.3 dapat dilihat sebagai berikut: [4]

$$P(H|E_1, E_2, E_3) = \frac{P(E_1|H) * P(E_2|H) * P(E_3|H) * P(H)}{P(E_1) * P(E_2) * P(E_3)} (2.3)$$

b. Naïve Bayes untuk Klasifikasi

Kaitan antara *Naïve bayes* dengan klasifikasi, korelasi hipotesis, dan bukti dengan klasifikasi adalah bahwa hipotesis dalam teorema Bayes merupakan label kelas yang menjadi target pemetaan dalam klasifikasi, sedangkan bukti merupakan fitur-fitur yang menjadi masukan dalam model klasifikasi. Jika X adalah vektor masukan yang berisi fitur dan Y adalah label kelas, naïve bayes dituliskan dengan P(Y|X). Notasi tersebut berarti probabilitas label kelas Y didapatkan setelah fitur-fitur X diamati. Notasi ini disebut juga probabilitas akhir (*posterior probability*) untuk Y, sedangkan P(Y) disebut probabilitas awal (*prior probability*). Dengan membangun model tersebut, suatu data uji X' dapat diklasifikasikan dengan mencari nilai Y' dengan memaksimalkan nilai P(Y'|X') yang didapat. [9]

Formulasi Naïve Bayes untuk klasifikasi seperti dalam persamaan 2.4. [9]

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^q P(X_i|Y)}{P(X)} \dots \dots \dots (2.4)$$

Keterangan:

P(Y|X) adalah probabilitas data dengan vector X pada kelas Y. P(Y) adalah probabilitas awal kelas Y dari semua fitur dalam vector X. karena P(X) selalu tetap, sehingga dalam perhitungan prediksi nantinya cukup hanya dengan menghitung P(Y) $\prod_{i=1}^q P(X_i|Y)$ dengan memilih yang besar sebagai kelas yang dipilih sebagai hasil prediksi. Sementara probabilitas independen $\prod_{i=1}^q P(X_i|Y)$ tersebut merupakan pengaruh semua fitur dari data terhadap setiap kelas Y, yang dinotaskan dengan persamaan 2.5. [9]

$$P(X|Y=y) = \prod_{i=1}^q P(X_i|Y = y) \dots \dots \dots (2.5)$$

Setiap fitur X = {X₁, X₂, X₃,, X_q} terdiri atas q atribut (q dimensi).

Umumnya, metode Naïve Bayes ini mudah dihitung untuk fitur bertipe kategoris. Namun untuk tipe numerik (kontinu), ada perlakuan khusus sebelum dimasukkan dalam Naïve Bayes, yaitu:

1. Melakukan diskretisasi pada setiap fitur kontinu dan mengganti nilai fitur kontinu tersebut dengan nilai interval diskret. Pendekatan ini dilakukan dengan mentransformasi fitur kontinu kedalam fitur ordinal.
2. Dari distribusi probabilitas diasumsikan bentuk tertentu untuk fitur kontinu dan memperkirakan parameter distribusi dengan data peralihan. Distribusi Gaussian biasanya dipilih untuk mempresentasikan probabilitas bersyarat dari fitur kontinu pada sebuah kelas P(X_i|Y), sedangkan distribusi Gaussian dikarakteristikan dengan dua parameter mean μ dan varian σ². Untuk setiap kelas y_j, probabilitas bersyarat kelas y_j untuk fitur X_i adalah seperti persamaan 2.6. Persamaan 2.6 dapat sebagai berikut: [9]

$$P(X_i=x_i|Y=y_i) = \frac{1}{\sqrt{2\pi} \sigma_{ij}} \exp \frac{-(x_i - \mu_{ij})^2}{2\sigma_{ij}^2} \dots \dots \dots (2.6)$$

Keterangan:

Parameter μ_{ij} bisa di dapat dari mean sampel X_i(\bar{X}) dari semua data latih yang menjadi milik kelas y_j sedangkan σ_{ij}² dapat diperkirakan dari varian sampel (S²) dari data latih.

c. Algoritma Naïve Bayes Classifier

Algoritma klasifikasi *naïve bayes* dihitung sesuai dengan rumus *Naïve Bayes* P(Y) $\prod_{i=1}^q P(X_i|Y)$

, yang langkah – langkah perhitungannya dijelaskan sebagai berikut: [2]

1. Menghitung nilai probabilitas kelas berdasarkan data latih seperti pada persamaan 2.7. [2]

$$P(Y) = \frac{Xy}{X} \dots\dots\dots (2.7)$$

Keterangan:

X : Jumlah total data
 Xy : Nama kelas/ output

2. Menghitung nilai probabilitas tiap fitur berdasarkan data latih seperti pada persamaan 2.8. [2]

$$\prod_{i=1}^q P(X_i|Y) = \frac{Xyx}{\sum X \in} \dots\dots\dots (2.8)$$

Keterangan:

Xyx : Data fitur bernilai y dari kelas X
 $\sum X \in$: Jumlah kelas X

3. Namun apabila data numerik maka:
 - a. Cari nilai mean dan standar deviasi dari masing – masing parameter yang merupakan dara numerik.

Adapun persamaan yang digunakan untuk menghitung nilai rata – rata hitung (mean) dapat dilihat pada persamaan 2.9 sebagai berikut: [2]

$$\mu = \frac{\sum_{i=1}^n X_i}{n} \dots\dots\dots (2.9)$$

atau pada persamaan 2.10 sebagai berikut:

$$\mu = \frac{x_1+x_2+x_3+\dots+x_n}{n} \dots\dots\dots (2.10)$$

dimana:

μ = rata – rata hitung (mean)
 x_i = nilai sample ke – i
 n = jumlah sample

Dan persamaan untuk menghitung nilai simpangan baku (standar deviasi) dapat dilihat pada persamaan 2.11 sebagai berikut: [2]

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}} \dots\dots\dots (2.11)$$

dimana:

σ = standar deviasi
 μ = rata – rata hitung (mean)
 x_i = nilai x ke – i
 n = jumlah sample

- b. Menghitung fitur bernilai numerik menggunakan rumus seperti persamaan 2.12. [2]

$$P(X_i=x_i|Y=y_i) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} \exp \frac{-(x_i - \mu_{ij})^2}{2\sigma_{ij}^2} \quad (2.12)$$

Fitur numerik berikut ini dihitung tiap data uji.

4. Menghitung nilai probabilitas akhir.

Mengalikan hasil dari probabilitas awal dan probabilitas setiap fitur atau P(Y) dan $\prod_{i=1}^q P(X_i|Y)$ pada masing – masing kelas data uji.

5. Data uji akan diklarifikasikan pada kelas dengan nilai probabilitas akhir terbesar.

3. METODE PENELITIAN

3.1 Pengumpulan Data

Pengumpulan data adalah suatu metode dan prosedur yang harus digunakan untuk mendapatkan sebuah informasi tentang apa saja yang harus dilakukan dalam untuk mengimplementasikan *Metode Naïve Bayes Classifier (NBC)* kedalam proses pengklasifikasian pemilihan penerima beras miskin (raskin) berdasarkan kriteria, atribut, data sampel, dan data pengujian. Pada tahap pengumpulan data ini terdapat beberapa hal yang harus dilakukan untuk membangun sebuah sistem, diantaranya adalah sebagai berikut:

- 1) Observasi
 Observasi yaitu suatu kegiatan dengan melakukan pengamatan pada suatu objek atau bidang yang sedang diteliti untuk mendapatkan informasi – informasi tambahan yang terkait dengan penelitian ini. Pengamatan ini dilakukan dengan melakuka riset untuk mengamati secara langsung aktivitas atau prosedur penyaluran dan pembagian beras miskin yang sedang berjalan di Kelurahan Desa Tambakmerang.
- 2) Wawancara
 Wawancara adalah kegiatan yang dilakukan dalam rangka mengumpulkan informasi dan data yang diperlukan untuk membangun sebuah sistem dengan melakukan tatap muka dan tanya jawab langsung dengan orang yang ahli dibidang yang akan diteliti. Dalam penelitian ini, peneliti melakukan wawancara pada beberapa sumber yang terkait dengan penelitian. Beberapa diantaranya adalah Kepala Kelurahan Desa Tambakmerang Bapak Margono, Kepala Seksi Kesejahteraan Kelurahan Desa Tambakmerang Bapak Waryatno. Peneliti mendapatkan data – data berupa *soft file* dan *hard copy* yang kemudian dipindahkan kedalam *Microsoft Excel* dimana data tersebut akan dibutuhkan dalam penelitian, untuk dijadikan kriteria, atribut, data sampel dan data pengujian dala pengolahan data dalam menentukan klasifikasi pemilihan penerima beras miskin.

Data yang diperoleh berupa 120 data yang nantinya 100 data untuk data sampel dan 20 data pengujian. 100 data sampel dimana 60 data klasifikasi layak menerima raskin dan 40 data tidak layak menerima raskin.

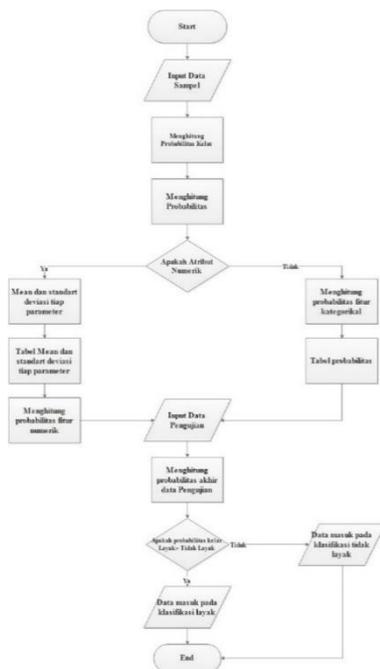
- 3) Metode Pustaka
Memperoleh data dengan cara membaca, mempelajari buku – buku, jurnal hasil penelitian dan literatur yang berhubungan dengan metode dan kasus yang berkenaan.

3.2 Pengambilan Data

Pada penelitian ini, data yang diambil adalah data sampel, data pengujian serta data kriteria dan *attribute* yang digunakan untuk menentukan penerima beras miskin. Data ini diperoleh dari hasil observasi dan wawancara.

3.3 Analisis Perancangan

Pada tahap analisis dan perancangan ini adalah tahap yang menspesifikasi bagaimana sistem dalam memenuhi kebutuhan informasi. Untuk dapat memenuhi kebutuhan pengguna, sistem ini akan memerlukan beberapa tahap desain seperti desain input, desain output, desain basis data, desain proses dan desain interface, selain itu pada desain sistem nanti akan diberikan gambaran secara detail tentang DAD dan ERD pada sistem. Gambar 3.1 berikut ini adalah flowchart sistem:

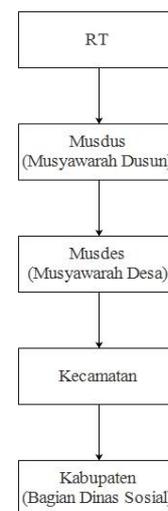


Gambar 3. 1 Flowchart Sistem

4. ANALISA DAN PERANCANGAN SISTEM

4.1 Analisa Sistem

Setelah melakukan observasi dan wawancara di Kelurahan Desa Tambakmerang dapat diketahui bahwa kegiatan penerimaan beras miskin di kelurahan ini tidak mengacu pada Pedoman Pelaksanaan Bantuan Pangan Non Tunai Tahun 2017. Pada pedoman ini penerima manfaat beras miskin adalah keluarga, yang selanjutnya disebut Keluarga Penerima Manfaat (KPM), dengan kondisi social ekonomi 25% terendah di daerah pelaksanaan. Dalam penentuan calon penerima beras miskin (Raskin) selama ini dilakukan oleh seksi kesejahteraan dengan usulan dari setiap dusun. Setiap dusun yang diberikan tugas untuk mengusulkan calon penerima raskin tidak mengacu pada kriteria keluarga miskin yang telah ditetapkan. Kriteria – kriteria yang ditentukan hanya berdasarkan penilaian kira – kira atau subjektif saja, sehingga banyak menimbulkan kecurangan atau tidak tepat sasaran dalam pembagian beras miskin (Raskin). Dengan menganalisis sistem yang sedang berjalan mempermudah mengevaluasi permasalahan dan kelemahan pada sistem agar sistem yang akan dibangun dapat sesuai dengan kebutuhan dan permasalahan yang ada. Untuk mengetahui alur sistem yang sedang berjalan maka diperlukan adanya urutan alur pembagian raskin di kelurahan. Adapun alur pembagian raskin mulai dari awal hingga ke penerima dapat dilihat pada Gambar 4.1. berikut ini:

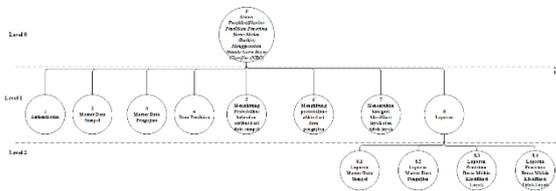


Gambar 4.1 Alur Penyaluran Beras Miskin Kelurahan Desa Tambakmerang

4.2 Rancangan Sistem

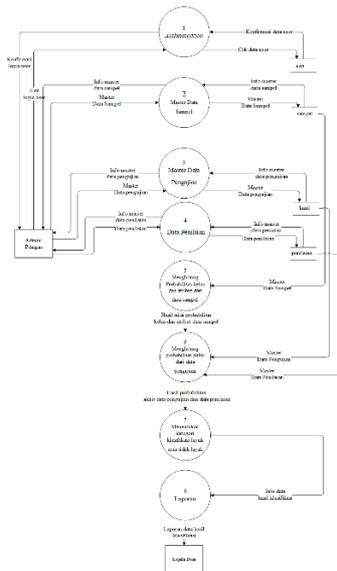
Untuk lebih memahami aliran sistem yang berjalan, maka diperlukan suatu gambaran sistem dan arus data yang digunakan mulai dari proses input hingga proses output yang akan dibutuhkan oleh pengguna. Berikut ini adalah gambaran sistem dan arus data yang dimodelkan kedalam DFD level 0 (Diagram Konteks), Diagram Jenjang, DFD level 1, dan DFD level 2.

- Diagram jenjang menerangkan kegiatan atau proses – proses yang akan dilalui pada sistem pengklasifikasian pemilihan penerima beras miskin. Diagram jenjang pada Sistem Pengklasifikasian Pemilihan Penerima Beras Miskin (Raskin) Menggunakan Metode *Naïve Bayes Classifier (NBC)* dapat dilihat pada Gambar 4.2.



Gambar 4.2 Diagram Jenjang

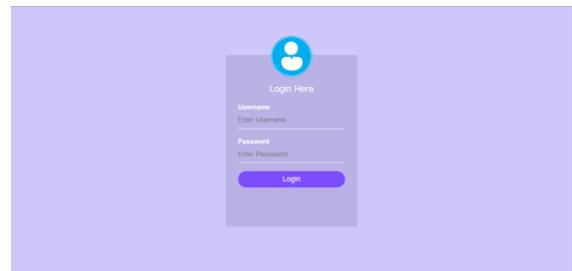
- Data Flow Diagram (DFD) Level 1* menjelaskan tentang gambaran proses yang lebih rinci dari diagram konteks dan menjelaskan proses sebuah penyimpanan. *Data Flow Diagram (DFD) Level 1* pada Sistem Pengklasifikasian Pemilihan Penerima Beras Miskin (Raskin) Menggunakan Metode *Naïve Bayes Classifier (NBC)* dapat dilihat pada Gambar 4.3.



Gambar 4.3 DFD Level 1

5. IMPLEMENTASI SISTEM

Pada bab ini akan membahas tentang tahap implementasi dan pembahasan program aplikasi dari sistem pengklasifikasi pemilihan penerima beras miskin menggunakan metode *Naïve Bayes Classifier*. Sistem ini bertujuan untuk menjabarkan kebutuhan yang diperlukan agar mudah dimengerti. Sistem pengklasifikasian pemilihan penerima beras miskin dirancang untuk dapat menilai penerima beras miskin, apakah layak mendapatkan atau tidak layak mendapatkan. Implementasi pada tahap ini bertujuan untuk menterjemahkan keperluan perangkat lunak ke dalam bentuk sebenarnya yang dimengerti oleh komputer atau dengan kata lain tahap implementasi ini merupakan tahapan lanjutan dari tahap perancangan yang sudah dilakukan. Dalam tahap implementasi ini akan dijelaskan mengenai perangkat keras (hardware) dan perangkat lunak (software) yang digunakan dalam membangun sistem ini, file-file yang digunakan dalam membangun sistem, implementasi pengujian sistem, tampilan web beserta potongan-potongan script program untuk menampilkan halaman web.



Gambar 5. 1 Halaman Login

Halaman login adalah halaman yang pertama kali akan muncul saat aplikasi Sistem Pengklasifikasian Pemilihan Penerima Beras Miskin (Raskin) Menggunakan Metode *Naïve Bayes Classifier (NBC)* diakses. Halaman ini berekstensi *.php* dengan nama file *index.php*. Untuk admin/petugas dapat masuk ke menu utama maka harus memasukkan username dan password yang sebelumnya sudah dibuat terlebih dahulu. Halaman login ditunjukkan pada Gambar 5.1

Gambar 5. 2 Halaman Data Master Sampel

Gambar 5.2 menunjukkan Halaman data master sampel yang merupakan halaman yang digunakan untuk menampilkan data sampel. Data sampel nantinya digunakan dalam pembuatan model probabilitas atau perhitungan klasifikasi pemilihan penerima beras miskin. Data dapat diimport dari bentuk *.xls*.

Gambar 5. 3 Halaman Data Master Pengujian

Gambar 5.3 menunjukkan Halaman data master pengujian yang digunakan untuk menampilkan data pengujian. Data pengujian merupakan data yang digunakan untuk menguji model hasil probabilitas metode *Naïve Bayes Classifier*. Dalam halaman ini terdapat hasil akurasi dan error dari model yang digunakan. Data dapat diubah dengan mengimport dari bentuk *.xls*. Hasil pengujian menggunakan model ini menghasilkan klasifikasi NBC atau klasifikasi dari sistem, jika klasifikasi dari data asli tidak sama dengan klasifikasi NBC maka data akan berwarna merah.

Gambar 5. 4 Halaman Laporan Data Pengujian

PEMERINTAHAN KABUPATEN WONOGIRI
KECAMATAN GIRIMARTO
SEKRETARIAT DESA TAMBAKMERANG
(Alamat: E. Street No.22, Telp. 70-704)

DATA PENGUJIAN

No	Nama Penduduk	NIK	Alamat	Klasifikasi NBC	Klasifikasi Awal
1	MIFTI	311221407230015	TAMBAKMERANG RT 01 RW 08	TIDAK LAYAK	LAYAK
2	LAINO	3112220402710001	TAMBAKMERANG RT 01 RW 08	TIDAK LAYAK	TIDAK LAYAK
3	SUBONO	311222104460001	TAMBAKMERANG RT 01 RW 08	LAYAK	TIDAK LAYAK
4	KARDE	311222142400001	TAMBAKMERANG RT 01 RW 08	TIDAK LAYAK	TIDAK LAYAK
5	SUTIMIN	311222010750004	TAMBAKMERANG RT 02 RW 08	TIDAK LAYAK	LAYAK
6	TARSO SEMITO	311222041178001	TAMBAKMERANG RT 02 RW 08	TIDAK LAYAK	LAYAK
7	TIYARSO	311221806850007	TAMBAKMERANG RT 02 RW 08	TIDAK LAYAK	TIDAK LAYAK
8	MARIMIN KAROSEMITO	3112220805550001	GENENG RT 01 RW 09	LAYAK	LAYAK
9	TIMIN MINTO WYONO	3112220509540001	GENENG RT 01 RW 09	LAYAK	LAYAK
10	SADIMO	3112223112470008	GENENG RT 01 RW 09	LAYAK	LAYAK
11	PARGIVANTO	311221309810001	GENENG RT 01 RW 09	LAYAK	TIDAK LAYAK
12	SRI RAHAYU	3112224510720001	GENENG RT 02 RW 09	LAYAK	LAYAK
13	SAMIDIN KROMO SEMITO	311221212330001	GENENG RT 03 RW 09	LAYAK	LAYAK
14	HARNI	311222409530003	GENENG RT 03 RW 09	TIDAK LAYAK	TIDAK LAYAK
15	KASIRI	3112220107280006	GENENG RT 04 RW 09	LAYAK	LAYAK

Halaman 10

Gambar 5. 5 Hasil Laporan Data Pengujian

Gambar 5.4 menunjukkan Halaman laporan data pengujian berisi data – data dari master data pengujian. Terdapat button cetak untuk mencetak laporan. Hasil dari mencetak laporan berbentuk pdf yang ditunjukkan pada Gambar 5.5.

6. PENUTUP

6.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan serta pembahasan, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Telah dibangun sistem pengklasifikasian pemilihan penerima beras miskin (raskin) menggunakan metode *Naïve Bayes Classifier (NBC)* di Kelurahan Tambakmerang, Girimarto, Wonogiri. Dibuat berbasis website sehingga dapat diakses Admin/Petugas dimanapun dan kapanpun. Berisi halaman master data sampel, model metode *Naïve Bayes Classifier (NBC)*, data pengujian, seleksi serta laporan. Dalam seleksi, Admin/ Petugas dapat menyeleksi data tunggal (satu data) atau data kelompok (banyak data). Aplikasi dibangun menggunakan bahasa pemrograman *PHP* dengan text editor *Sublime Text 3* dan engine *SQLyog*.
2. Pengujian metode *Naïve Bayes Classifier (NBC)* untuk pengklasifikasian pemilihan penerima beras miskin (raskin) menggunakan 100 data sampel dan 20 data pengujian menghasilkan akurasi modelnya sebesar 70 % dan error sebesar 30 %.
3. Perbandingan hasil akurasi metode *Naïve Bayes Classifier (NBC)* antara sistem yang dibangun dengan Aplikasi *RapidMiner Studio 9.2* menghasilkan akurasi yang sama, perbandingannya dapat dilihat pada Sub Bab 5.6. Setelah dilakukan beberapa kali uji coba didapatkan hasil akurasi terbesar yaitu menggunakan 100 data sampel dan 20 data

pengujian menghasilkan akurasi yang sama yaitu sebesar 70% dapat dilihat pada gambar 5.24 dan 5.25.

6.2 Saran

Adapun saran – saran yang dapat disampaikan untuk pengembang selanjutnya yang mungkin dapat membangun adalah sebagai berikut:

- a. Pada sistem ini hanya menggunakan 100 data sampel, 20 data pengujian, beberapa kriteria – kriteria, dan attribut dari calon pemilihan penerima beras miskin (raskin) yang bersifat statis sehingga tidak dapat diubah. Pada pengembangan sistem berikutnya data sampel, data pengujian dapat ditambah lebih banyak lagi dan bisa dibuat dinamis, begitu pula dengan kriteria – kriterianya. Semakin banyak data sampel maka sistem akan semakin detail dan akurat.
- b. Sistem ini masih terdapat kekurangan pada import data, yaitu belum mampu import data dalam bentuk selain *.xls* dan cetak dalam bentuk excel, sehingga untuk pengembangan selanjutnya dapat ditambahkan import dalam ekstensi *.xlsx* serta menu cetak dalam bentuk excel.
- c. Untuk pengembangan selanjutnya dapat mengimplemmentasikan menggunakan metode lain, yaitu penerapan Jaringan Saraf Tiruan. Dengan menggunakan data yang sudah tersedia sistem jaringan syaraf belajar dari berbagai pola masukan yang tersedia di dataset dan menyesuaikan koneksi bobot untuk mencapai hasil yang diharapkan sehingga nantinya akan menghasilkan tingkat akurasi yang lebih tinggi.

UCAPAN PERSEMBAHAN

Naskah Publikasi ini dapat diselesaikan tidak lepas dari segala bantuan, bimbingan, dorongan dan doa dari berbagai pihak, yang pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Kepada Bapak Agus Sujarwadi S.Kom., M.T. selaku Dosen Pembimbing yang telah banyak memberikan petunjuk dalam penyusunan naskah publikasi ini.
2. Teristimewa kepada Orang Tua penulis yang selalu mendoakan, memberikan motivasi dan pengorbanannya baik dari segi moril maupun materi kepada penulis sehingga penulis dapat menyelesaikan penelitian ini.

DAFTAR PUSTAKA

- [1]Adi, S. (2018), *Implementasi Algoritma Naive Bayes Classifier Untuk Klasifikasi Penerima Beasiswa PPA Di Universitas Amikom Indonesia, Jurnal Mantik Penusa*, 22(1), 11–16.
- [2]Alfa, S. (2015), *Implementasi Metode Klasifikasi Naïve Bayes dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga*, *Citec Journal*, 2(3), 207–217Diakses dari ojs.amikom.ac.id/index.php/citec/article/download/375/355.
- [3]Angrawati, D., Yamin, M. dan Ransi, N. (2016), *Sistem Pendukung Keputusan Menentukan Jumlah Beras Miskin Menggunakan Metode Simple Additive Weight (SAW)*, *semanTIK*, 2(1), 39–45.
- [4]Aries, F. (2015), *Aplikasi Penentuan Penerimaan Beras Miskin (Raskin) Di Ds. Sidomulyo Kec. Deket Kab. Lamongan Dengan Metode Naive Bayes*, .
- [5]Fadlan, C., Ningsih, S. dan Windarto, A.P. (2018), *Penerapan Metode Naive Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra*, *Jurnal Teknik Informatika Musirawas (JUTIM)*, 3(1), 1.
- [6]Kurniawan, I. dan Saputra, R.A. (2017), *Penerapan Algoritma C5 . 0 Pada Sistem Pendukung Keputusan Kelayakan Penerimaan Beras Masyarakat Miskin*, *Jurnal Informatika*, 4(2), 236–240.
- [7]Lorosae, T.A., Prakoso, B.D., Saifudin dan Kusri (2018), *Analisis Sentimen Berdasarkan Opini Masyarakat Pada Twitter Menggunakan Naive Bayes*, *Seminar Nasional Teknologi Informasi dan Multimedia*, 25–30.
- [8]Nofriansyah, D. (2014), *Konsep Data Mining VS Sistem Pendukung Keputusan*, ed. 1 Yogyakarta: Deepublish.
- [9]Prasetyo, E. (2012), *Data Mining Konsep dan Aplikasi Menggunakan Matlab*, Yogyakarta: Andi.
- [10]Vulandari, R.T. (2017), *Data Mining Teori dan Aplikasi Rapidminer*, ed. 1 Yogyakarta: Penerbit Gava Media.
- [11]Wibowo, A.P. dan Hartati, S. (2016), *Sistem Klasifikasi Kinerja Satpam Menggunakan Metode Naïve Bayes Classifier*, *Jurnal Invotek Polbeng-Seri Informatika*, 1(2), 192–201.