

Optimasi Naïve Bayes Dan Algoritma Genetika Untuk Prediksi Penerimaan Beasiswa Pendidikan Pada SMP Utama

Nining Suryani¹, Evy Priyanti²

¹Program Studi Sistem Informasi Kampus Kabupaten Karawang
Universitas Bina Sarana Informatika
Jl. Kamal Raya No. 18 Ringroad Barat, Cengkareng, Jakarta Barat
¹nining.nns@bsi.ac.id

²Program Studi Komputerisasi Akuntansi
Universitas Bina Sarana Informatika
Jl. Kamal Raya No. 18 Ringroad Barat, Cengkareng, Jakarta Barat
²evy.evp@bsi.ac.id

Abstract - - Educational scholarships are one of the efforts to sustain students in getting a better education. Not a few students drop out in the middle of the road or cannot continue their education at the same level or higher level. Selection according to the criteria for scholarship recipients is important so that scholarships are on target. Similar to Depok Primary Middle School, educational scholarships are provided by schools based on 9 criteria for scholarship recipients, namely parent status, parent work, rented house, home appliances, vehicles, parents' savings, parents' jewelry, cellphones and pocket money. With the number of prospective scholarship recipients there is an algorithm needed to accurately predict students who are entitled to scholarships. With the naive bayes algorithm, accuracy is 77.50% in predicting scholarship recipients based on the criteria found in students. The use of genetic algorithms is done to get a more optimal level of accuracy. This is evidenced by the accuracy of 83.33%.

Keywords: optimization, naive bayes, acceptance, genetic algorithm

PENDAHULUAN

Angka putus sekolah pada jenjang SMP terus naik setiap tahunnya, berdasarkan data Kemendikbud yang menyatakan data putus sekolah tahun ajaran 2017/2018 berjumlah 51190. Angka ini meningkat 24,39% dibandingkan tahun sebelumnya dan propinsi Jawa Barat memiliki angka putus sekolah terbanyak yaitu 9340 pada tahun ajaran 2017/2018. Angka ini meningkat 7,54% dibandingkan tahun sebelumnya. Putus sekolah didasari oleh banyak faktor, salah satunya adalah faktor kemiskinan. Kemiskinan ini kemudian menjadi alasan seseorang mengajukan permohonan beasiswa kepada lembaga pendidikan agar tidak putus sekolah. Berbagai langkah pun ditempuh, antara lain melalui ketersediaan dana pembebasan biaya sekolah melalui program Surat Keterangan Tidak Mampu (SKTM) yang dikeluarkan dari kantor kecamatan setempat. Adakalanya memang benar sang pelajar tidak mampu secara finansial, tapi tak jarang dari kalangan mampu pun memanfaatkan surat ini. Beasiswa diartikan sebagai bentuk penghargaan yang diberikan kepada individu agar dapat melanjutkan pendidikan ke jenjang yang lebih tinggi. Penghargaan itu dapat berupa akses tertentu pada suatu instansi atau penghargaan berupa bantuan keuangan (Murniasih, 2008). Beasiswa merupakan pemberian bantuan keuangan yang diberikan perorangan, mahasiswa atau pelajar yang digunakan demi kelangsungan pendidikan yang ditempuh

(Tullah, Mariana, & Baskoro, 2018)

Beberapa katagori beasiswa berdasarkan (Murniasih, 2009) adalah

1. Beasiswa Bantuan
Beasiswa bantuan adalah beasiswa untuk mendanai kegiatan akademik para mahasiswa yang kurang beruntung, tetapi memiliki prestasi.
2. Beasiswa Atletik
Beasiswa yang diberikan pada cabang olah raga khusus dibidang atletik.
3. Beasiswa penghargaan, biasanya diberikan kepada kandidat yang memiliki keunggulan akademik. Beasiswa ini diberikan berdasarkan prestasi akademik mereka secara keseluruhan.
4. Banyak orang menilai bahwa beasiswa diberikan kepada penerimanya untuk menutupi keperluan akademik secara keseluruhan atau biasa disebut beasiswa penuh
5. Bantuan yang diberikan kepada mahasiswa yang mengalami kesulitan membayar biaya pendidikannya adalah berupa Beasiswa Bantuan Belajar Mahasiswa (BBM), sama dengan PPA tujuannya membantu meringankan beban orang tua dari kalangan ekonomi lemah.
6. Beasiswa Bidik Misi
Beasiswa bidik merupakan program dari Dirjen Dikti untuk perguruan tinggi negeri. Beasiswa bagi

calon mahasiswa yang kurang mampu tapi berpotensi.

7. Beasiswa Peningkatan Pretasi Akademik (PPA)

Beasiswa Peningkatan Pretasi Akademik adalah beasiswa yang diberikan untuk peningkatan pemerataan dan kesempatan belajar bagi siswa/mahasiswa yang mengalami kesulitan membayar biaya pendidikan sebagai akibat krisis ekonomi, terutama bagi mahasiswa/siswa yang berprestasi akademik.

Adapun tujuan beasiswa peningkatan prestasi akademik secara umum, yaitu :

- Meningkatkan pemerataan dan kesempatan belajar bagi mahasiswa/siswa yang mengalami kesulitan membayar pendidikan.
- Mendorong dan mempertahankan semangat belajar mahasiswa agar mereka dapat menyelesaikan studi/ pendidikan tepat waktunya.
- Mendorong untuk meningkatkan prestasi akademik sehingga memacu peningkatan kualitas pendidikan.

SMP Utama didirikan tahun 2003 dengan semangat idealisme pengabdian untuk bangsa oleh Yayasan Amal Khair Yasmin, Tangerang Selatan. Sekolah memiliki jumlah siswa yang diterima berdasarkan daya tampung sarana prasarana dan fasilitas yang dimiliki, serta jumlah tenaga pendidik dan kependidikan, sebanyak 30 – 32 siswa perkelas setiap tahun. Seleksi penerima beasiswa yang dilaksanakan sekolah masih dilakukan secara konvensional sehingga cenderung tidak efisien dan hasil seleksi kurang akurat yang dapat menyebabkan rentan terjadi kesalahan sehingga tidak tepat sasaran penerima beasiswa sekolah. Untuk itu diperlukan penggunaan metode untuk menentukan tingkat akurasi yang tinggi untuk mengklasifikasi penerima beasiswa berdasarkan kriteria sekolah agar mendapatkan penerima beasiswa yang benar-benar membutuhkan.

METODOLOGI PENELITIAN

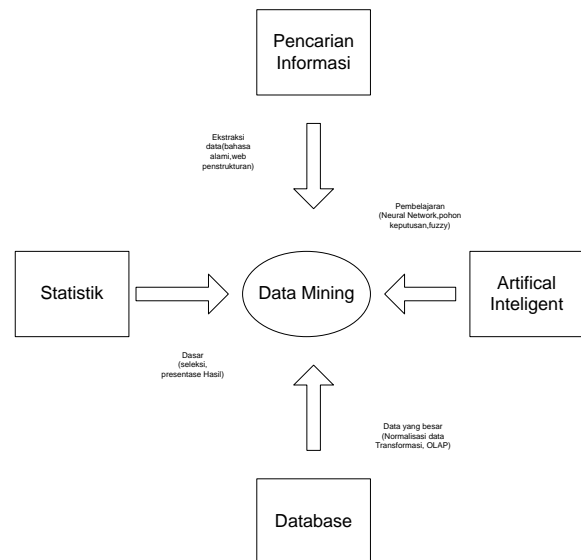
A. Data Mining

Berdasarkan (Prasetyo, 2014) pengertian data mining cukup sulit dijelaskan dengan gambar jika mengingat data mining juga merupakan gabungan dari beberapa bidang ilmu.

Sedangkan, data mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistic, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar (Kusrini & Emha Taufiq Luthfi, 2009) .

Karena, data mining berkaitan dengan pengumpulan data dari database yang besar, maka data mining juga dikenal dengan nama knowledge-discovery in database (KDD) (Prabowo Pudjo Widodo, Rahmadya Trias Handayanto, 2013)

Walaupun istilah KDD dan data mining sebenarnya memiliki konsep yang berbeda, karena data mining merupakan salah satu tahapan dari keseluruhan proses KDD (Kusrini & Emha Taufiq Luthfi, 2009). Berdasarkan definisi tersebut, data mining merupakan gabungan dari beberapa bidang ilmu seperti yang digambarkan pada gambar dibawah ini:



Gambar 1. Proses Data Mining
Sumber : (Kusrini & Emha Taufiq Luthfi, 2009)

Data mining dapat dibedakan menjadi 6 kelompok menurut (Kusrini & Emha Taufiq Luthfi, 2009) berdasarkan tugas yang dapat dilakukan, yaitu:

- Deskripsi
- Estimasi
- Prediksi
- Klasifikasi
- Pengklusteran
- Asosiasi

B. Naive Bayes

Klasifikasi adalah salah satu tugas yang penting dalam data mining, dalam klasifikasi sebuah pengklasifikasi dibuat dari sekumpulan data latih dengan kelas yang telah di tentukan sebelumnya. Performa pengklasifikasi biasanya diukur dengan ketepatan (atau tingkat galat) (Prasetyo, 2012) .

Metode ini diperkenalkan oleh Thomas Bayes untuk memprediksi kemungkinan yang akan terjadi di masa yang akan datang berdasarkan pengalaman di masa lalu (Meilani & Susanti, 2015) .

Metode ini menurut (Kusrini & Emha Taufiq Luthfi, 2009) “adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class”.

Bentuk umum dari naive bayes atau yang lebih dikenal dengan teorema bayes menurut (Kusrini & Emha Taufiq Luthfi, 2009) adalah sebagai berikut:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

Gambar 2
Persamaan umum teorema bayes
Sumber : (Kusrini & Emha Taufiq Luthfi, 2009)

Penjelasan dari teorema bayes ini menurut (Kusrini & Emha Taufiq Luthfi, 2009) adalah sebagai berikut:

X= Data dengan class yang belum diketahui
H= Hipotesis data X merupakan suatu class spesifik
P(H|X)= Kemungkinan hipotesis H berdasarkan kondisi X (Posterior Probability)
P(H)= Kemungkinan hipotesis H (Prior Probability)
P(X|H)= Kemungkinan kondisi X berdasarkan kondisi hipotesis H
P(X)= Kemungkinan dari kondisi X
Teorema bayes dapat diterapkan dalam klasifikasi yang lebih dikenal dengan naïve bayes classification dengan bentuk umum menurut (Santosa, 2007) adalah:

$$v_{MAP} = \text{arg max}_{v_j \in V} P(a_1, a_2, a_3, \dots, a_n | v_j) P(v_j)$$

Gambar 3
Penerapan teorema bayes dalam klasifikasi
Sumber : (Santosa, 2007)

Apabila persamaan tersebut didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output, maka penerapan tersebut akan menjadi sebuah persamaan baru (Santosa, 2007), yaitu:

$$v_{NB} = \text{arg max}_{v_j \in V} P(v_j) \prod P(a_i | v_j)$$

Gambar 4
Persamaan naïve bayes classification
Sumber : (Santosa, 2007)

Dimana v_{NB} merupakan nilai output hasil klasifikasi naïve bayes (Santosa, 2007) .

C. Normalisasi Data

Sebelum dilakukan pelatihan pada dataset, data primer harus dinormalisasikan dahulu dengan tujuan agar proses pelatihan berjalan dengan cepat dan mampu digunakan untuk melakukan pelatihan. Persamaannya sebagai berikut :

$$x^1 = \frac{0.8(x - a)}{b - a} + 0.1$$

Keterangan : x^1 = Nilai Transform
X = Nilai Asli

a = Nilai Minimal
b = Nilai Maximal

Berdasarkan sampel acak nilai output (class) beasiswa Test sebagai berikut :

Data Beasiswa x=1

Diketahui :

Nilai Minimum (a) = 0

Nilai Maximum (b) = 1

Data yang akan ditransform (x)=0.9.

Dataset yang telah ditransformasikan dengan perhitungan fungsi aktivasi sigmoid biner untuk dataset pemberian beasiswa dapat dilihat pada tabel 1 Seperti contoh pada record pertama setiap atribut telah ditransformasi nilai awalnya menjadi nilai baru berturut-turut (0.9, 0.1,0.9,0.9,0.1,0.1,0.1,0.1,0.9).

Tabel 1 Kriteria Beasiswa

			R						C
			u						l
			m						s
			h						a
			,	per	ken	tab	perh	han	a
			,	ala	ken	ung	iasa	dph	n
			k	ken	an	an	an	one	g
			o	tan	dara	an	an	oran	g
			nt	ru	an,	kel	oran	oran	J
			ra	ma	an,	uar	g	g	a
			h	ah	ah	ga	tua	tua	j
			n,						a
			ta						n
			na						is
			h						w
									a
A									
h									
m									
ad	1	0	1	1	0	0	0	0	1
de									Y
ni									
el									

D. Algoritma Genetika

Algoritma Genetika atau Genetic Algorithm disingkat dengan GA merupakan algoritma yang memanfaatkan proses seleksi alamiah yang dikenal sebagai proses evolusi.

Algoritma Genetika memiliki istilah-istilah yang digunakan, dimana istilah-istilah tersebut mewakili elemen-elemen dalam teori yang dikemukakan oleh Charles Darwin, adapun istilah-istilah tersebut adalah (Obitko, 1998) :

a. Populasi

Sekumpulan solusi dari permasalahan yang akan diselesaikan menggunakan Algoritma Genetika. Population terdiri dari sekumpulan kromosom.

b. Kromosom

Mewakili sebuah solusi yang mungkin (feasible solution) untuk permasalahan yang ingin diselesaikan. Sebuah kromosom terdiri dari sekumpulan gen.

c. Gen

Mewakili elemen-elemen yang ada dalam sebuah solusi.

d. Parent

Kromosom yang akan dikenai operasi genetik (crossover).

e. Offspring

Kromosom yang merupakan hasil dari operasi genetik (crossover dan mutasi).

f. Crossover

Operasi genetik yang mewakili proses

perkembangbiakan antar individu. Dalam melakukan proses crossover dibutuhkan sepasang parent dan akan menghasilkan satu atau lebih keturunan (offspring).

g. Mutasi

Operasi genetik yang mewakili proses mutasi dalam perjalanan hidup individu. Peran mutasi adalah menghasilkan perubahan acak dalam populasi yang berguna untuk menambah variasi dari chromosome- chromosome dalam sebuah populasi.

h. Selection procedure

Proses yang mewakili seleksi alam (natural selection) dari teori Darwin. Proses ini dilakukan dalam memilih parent dari operasi genetik (crossover) yang akan dilakukan untuk menghasilkan keturunan (offspring).

i. Fitness value

Penilaian yang menentukan bagus tidaknya sebuah kromosom. Kromosom yang memiliki fitness value yang rendah pada akhirnya akan tersingkir oleh kromosom-kromosom yang memiliki fitness value yang lebih baik.

j. Evaluation function

Fungsi yang digunakan untuk menentukan nilai dari fitness value. Evaluation function ini merupakan sekumpulan kriteria-kriteria tertentu dari permasalahan yang ingin diselesaikan.

k. Generation

Satuan dari populasi setelah melalui operasi-operasi genetika, berkembang biak dan menghasilkan keturunan. Pada akhir dari setiap generation, untuk menjaga agar jumlah kromosom dalam populasi tetap konstan, maka kromosom yang memiliki fitness value yang rendah dan memiliki peringkat dibawah nilai minimal akan dihapus dari populasi.

Algoritma Genetika memiliki keunggulan-keunggulan dibandingkan dengan metode-metode heuristic yang lain, yaitu:

a. Algoritma Genetika menyelesaikan masalah dengan mengkodekan permasalahan menjadi kromosom, bukan dengan menyelesaikan permasalahan itu sendiri. Karena itu diperlukan pemodelan kromosom yang baik dan efektif yang dapat mewakili solusi dari permasalahan yang dihadapi.

b. Algoritma Genetika memulai prosesnya dengan sekumpulan initial solutions, berbeda dengan metaheuristic lain seperti Simulated Annealing dan Tabu Search yang memulai proses dengan sebuah solusi tunggal, dan berlanjut ke solusi lainnya melalui suatu transisi. Karena itu GA melakukan pencarian multi-directional dalam solution space, yang memperkecil kemungkinan berhentinya pencarian pada kondisi lokal optimum.

c. Hanya diperlukan sebuah fungsi evaluasi tunggal yang berbeda untuk tiap permasalahan.

d. Algoritma Genetika tidak mengetahui kapan dirinya telah mencapai solusi optimal.

HASIL DAN PEMBAHASAN

Dalam penelitian ini peneliti menggunakan 9 atribut / kriteria, yaitu atribut status orang tua, atribut

pekerjaan orang tua, atribut arribut kontrakan, atribut peralatan rumah, atribut kendaraan, atribut tabungan orang tua, atribut perhiasaan orang tua, atribut handphone, atribut uang jajan dengan sampel 200 data.

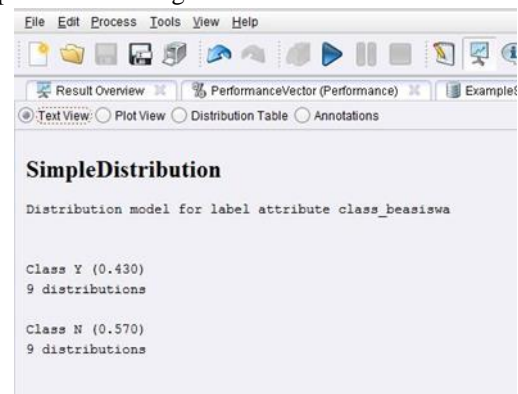
A. Penghitungan dengan algoritma naïve bayes Hasil dari perhitungan metode klasifikasi naïve bayes secara manual adalah sebagai berikut:

$$1. P(X|class_beasiswa="Yes")P(class_beasiswa = "Yes") = 0.018 \times 0.430 = 0.007$$

$$2. P(X|class_beasiswa="No")P(class_beasiswa = "No") = 0.003 \times 0.570 = 0.001$$

Karena nilai hasil perhitungan $P(X|class_beasiswa = "Yes") P(class_beasiswa = "Yes")$ lebih besar dari $P(X|class_menerima_beasiswa = "No")P(class_menerima_beasiswa = "No")$, maka sampel data ke 199 pada data arsip instrument tinjauan lapangan merupakan sampel yang termasuk kedalam kelas yang layak menerima beasiswa.

Distribution model menjelaskan distribusi data – data yang telah diimpor pada proses penempatan kelas menggunakan metode naïve bayes sebagaimana yang diperlihatkan oleh gambar berikut ini :



Gambar 5. Model Distribusi Data Arsip Instrumen Tinjauan lapangan Berdasarkan Kelas yang Terbentuk

Pada gambar 5 tersebut, terdapat 2 kelas yang terbentuk berdsarkan hasil pemrosesan data dengan metode naïve bayes, yaitu kelas Y dan kelas N.

Kelas Y merupakan kelas penerima beasiswa berdasarkan kriteria – kriteria yang terdapat pada peserta didik, sedangkan kelas N merupakan kelas yang tidak menerima beasiswa berdsarkan kriteria – kriteria yang tidak terdapat pada peserta didik. Setiap kelas, mempunyai 9 distribusi yang merupakan kriteria – kriteria yang telah diinput berdasarkan data arsip instrument tinjauan lapangan yang telah diimpor sebelumnya.

Angka yang ditunjukkan oleh setiap kelas, merupakan hasil dari perhiutngan teorema bayes, yaitu prior probability dari setiap kelas. Prior probability pada kelas yang menerima beasiswa di dapat dengan membagi jumlah yang menerima beasiswa, yaitu : 86 data dengan jumlah keseluruhan data yang ada, yaitu : 200 data. Begitu juga dengan

prior probability dari kelas yang tidak menerima beasiswa didapat dengan membagi jumlah data yang tidak menerima beasiswa, yaitu 114 data dengan jumlah keseluruhan yang ada, yaitu 200 data. Perhitungan masing – masing kelas tersebut adalah :

$$P(\text{class_beasiswa} = \text{"Y"}) = 86/200 = 0,430.$$

$$P(\text{class_beasiswa} = \text{"N"}) = 114/200 = 0,570$$

Setelah mendapatkan informasi tentang distribusi kriteria–kriteria yang dimasukan dalam kelas tertentu, maka distribution model menjelaskan bagaimana kedua kelas tersebut dapat terbentuk berdasarkan kriteria–kriteria yang telah diinput pada data arsip instrument tinjauan lapangan sebagaimana yang diperlihatkan oleh gambar berikut :

Attribute	Parameter	N	Y
statusortu	value=1	0.851	0.558
statusortu	value=0	0.149	0.442
statusortu	value=unknc	0.000	0.000
pekerjaan	value=0	0.439	0.465
pekerjaan	value=1	0.561	0.535
pekerjaan	value=unknc	0.000	0.000
rumah	value=1	0.561	0.372
rumah	value=0	0.439	0.628
rumah	value=unknc	0.000	0.000
peralatan	value=1	0.544	0.256
peralatan	value=0	0.456	0.744
peralatan	value=unknc	0.000	0.000
kendaraan	value=1	0.851	0.628
kendaraan	value=0	0.149	0.372
kendaraan	value=unknc	0.000	0.000
tabungan	value=0	0.588	0.546
tabungan	value=1	0.412	0.453
tabungan	value=unknc	0.000	0.000
perhiasaan	value=1	0.368	0.233
perhiasaan	value=0	0.632	0.767
perhiasaan	value=unknc	0.000	0.000
handphone	value=1	0.500	0.314
handphone	value=0	0.500	0.686
handphone	value=unknc	0.000	0.000
uangjajan	value=1	0.684	0.128
uangjajan	value=0	0.316	0.872
uangjajan	value=unknc	0.000	0.000

Gambar 6. Hasil Distribusi Kriteria – Kriteria pada Kelas Y dan Kelas N

Pada tabel distribusi kriteria – kriteria diatas, terdapat angka – angka yang merupakan hasil dari perhitungan berdasarkan metode naïve bayes.

Hasil tersebut akan digunakan untuk menarik kesimpulan apakah suatu sampel tertentu yang merupakan kriteria – kriteria penerima beasiswa yang terdapat pada peserta didik, termasuk yang menerima ataupun yang tidak menerima beasiswa. Berikut penjelasan dari atribut penerima beasiswa yaitu atribut status orang tua, atribut pekerjaan orang tua, atribut aribut kontrakan, atribut peralatan rumah, atribut kendaraan, atribut tabungan orang tua, atribut perhiasaan orang tua, atribut handphone, atribut uang jajan.

Algoritma naïve bayes dapat memprediksi secara cepat dan akurat data penerima beasiswa dengan perhitungan sebagai berikut :

$$\text{Posterior} = \text{prior} \times \text{likelihood evidence}$$

$$P(\text{status} = 1 \mid \text{class_menerima} = \text{Y}) = 48/86 = 0.558$$

$$P(\text{status} = 1 \mid \text{class_menerima} = \text{N}) = 97/114 = 0.850$$

$$P(\text{status} = 0 \mid \text{class_menerima} = \text{Y}) = 38/86 = 0.441$$

$$P(\text{status} = 0 \mid \text{class_menerima} = \text{N}) = 17/114 = 0.149$$

$$P(\text{Pekerjaan} = 1 \mid \text{class_beasiswa} = \text{Y}) = 46/86 = 0.534$$

$$P(\text{Pekerjaan} = 1 \mid \text{class_beasiswa} = \text{N}) = 64/114 = 0.561$$

$$P(\text{Pekerjaan} = 0 \mid \text{class_beasiswa} = \text{Y}) = 40/86 = 0.465$$

$$P(\text{Pekerjaan} = 0 \mid \text{class_beasiswa} = \text{N}) = 50/114 = 0.438$$

$$P(\text{rumah} = 1 \mid \text{class_menerima} = \text{Y}) = 32/86 = 0.372$$

$$P(\text{rumah} = 1 \mid \text{class_menerima} = \text{N}) = 54/114 = 0.473$$

$$P(\text{rumah} = 0 \mid \text{class_menerima} = \text{Y}) = 64/86 = 0.744$$

$$P(\text{rumah} = 0 \mid \text{class_menerima} = \text{N}) = 50/114 = 0.438$$

$$P(\text{peralatan} = 1 \mid \text{class_menerima} = \text{Y}) = 22/86 = 0.255$$

$$P(\text{peralatan} = 1 \mid \text{class_menerima} = \text{N}) = 62/114 = 0.543$$

$$P(\text{peralatan} = 0 \mid \text{class_menerima} = \text{Y}) = 64/86 = 0.744$$

$$P(\text{peralatan} = 0 \mid \text{class_menerima} = \text{N}) = 52/114 = 0.456$$

$$P(\text{kendaraan} = 1 \mid \text{class_menerima} = \text{Y}) = 54/86 = 0.627$$

$$P(\text{kendaraan} = 1 \mid \text{class_menerima} = \text{N}) = 97/114 = 0.850$$

$$P(\text{kendaraan} = 0 \mid \text{class_menerima} = \text{Y}) = 32/86 = 0.372$$

$$P(\text{kendaraan} = 0 \mid \text{class_menerima} = \text{N}) = 17/114 = 0.149$$

$$P(\text{tabungan} = 1 \mid \text{class_menerima} = \text{Y}) = 39/86 = 0.453$$

$$P(\text{tabungan} = 1 \mid \text{class_menerima} = \text{N}) = 47/114 = 0.412$$

$$P(\text{tabungan} = 0 \mid \text{class_menerima} = \text{Y}) = 47/86 = 0.546$$

$$P(\text{tabungan} = 0 \mid \text{class_menerima} = \text{N}) = 67/114 = 0.587$$

$$P(\text{perhiasaan} = 1 \mid \text{class_menerima} = \text{Y}) = 20/86 = 0.232$$

$$P(\text{perhiasaan} = 1 \mid \text{class_menerima} = \text{N}) = 66/114 = 0.578$$

$$P(\text{perhiasaan} = 0 \mid \text{class_menerima} = \text{Y}) = 42/86 = 0.488$$

$$P(\text{perhiasaan} = 0 \mid \text{class_menerima} = \text{N}) = 72/114 = 0.631$$

$$P(\text{handphone} = 1 \mid \text{class_menerima} = \text{Y}) = 27/86 = 0.313$$

$$P(\text{handphone} = 1 \mid \text{class_menerima} = \text{N}) = 57/114 = 0.500$$

$$P(\text{handphone} = 0 \mid \text{class_menerima} = \text{Y}) = 59/86 = 0.686$$

$$P(\text{handphone} = 0 \mid \text{class_menerima} = \text{N}) = 57/114 = 0.500$$

$$P(\text{uangjajan} = 1 \mid \text{class_menerima} = \text{Y}) = 11/86 = 0.127$$

$$P(\text{uangjajan} = 1 \mid \text{class_menerima} = \text{N}) =$$

$$78/114 = 0.684$$

$$P(\text{uangjajan} = 0 \mid \text{class_menerima beasiswa} = Y) = 75/86 = 0.872$$

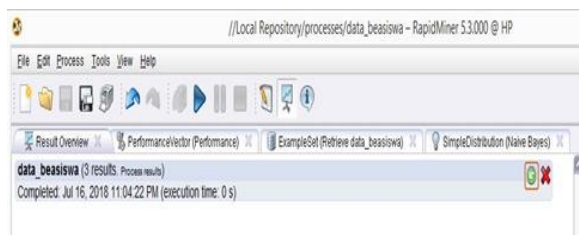
$$P(\text{uangjajans} = 0 \mid \text{class_menerima beasiswa} = N) = 36/114 = 0.315$$

Pada bagian ini, data akan diukur kecepatan pembacaan data dan akurasi ketepatan klasifikasi dengan menggunakan metode naïve bayes. Bagian ini merupakan bagian pengujian model. Setelah data diproses, maka hasil pengujian dari data arsip instrument tinjauan lapangan digambarkan sebagai berikut :

	true Y	true N	class precision
pred. Y	72	31	69.90%
pred. N	14	83	85.57%
class recall	83.72%	72.81%	

Gambar 7. Hasil Akurasi Pemrosesan data dengan Metode Naïve Bayes

Hal ini disebabkan adanya kelas yang tidak sesuai berdasarkan kriteria– kriteria yang telah diinput pada data arsip instrument tinjauan lapangan. Selain itu, metode naïve bayes memiliki kecepatan yang tinggi dalam mengklasifikasikan data – data yang telah diimport tersebut. Hal ini diperlihatkan oleh gambar berikut :



Gambar 8. Waktu yang dibutuhkan Untuk Mengklasifikasi Data

Pada gambar tersebut, waktu yang dibutuhkan oleh metode naïve bayes dalam mengklasifikasi data adalah 0 detik. Hal ini menunjukkan bahwa metode naïve bayes memiliki kecepatan yang tinggi dalam mengklasifikasi data.

B. Perhitungan Optimasi Model dengan Algoritma Genetika

Penelitian ini akan dilakukan perhitungan manual dalam pemilihan variabel menggunakan dataset yang ada. Perhitungan dilakukan dengan mengambil 5 record dari dataset beasiswa pada SMP Utama Depok.

Tahapan dalam metode Genetic Algorithm

berlangsung sebagai berikut :

- Menentukan Populasi Awal
C1: [0.9_0.1_0.9_0.9_0.1_0.1_0.1_0.9]
C2:[0.9_0.1_0.9_0.9_0.9_0.1_0.9_0.9]
C3:[0.9_0.9_0.9_0.9_0.9_0.1_0.1_0.9]
C4:[0.9_0.9_0.9_0.9_0.9_0.1_0.1_0.1]
C5:[0.1_0.9_0.9_0.9_0.9_0.1_0.1_0.9]

2. Evaluasi Nilai Fitness

$$\text{Fitness C1} = 1/1 + (0,9+0,1+0,9+0,9+0,1+0,1+0,1+0,9) - 1 = 6,500$$

$$\text{Fitness C2} = 1/1 + (0,9+0,1+0,9+0,9+0,9+0,1+0,9+0,9+0,9) - 1 = 6,500$$

$$\text{Fitness C3} = 1/1 + (0,9+0,9+0,9+0,9+0,9+0,1+0,1+0,9+0,9) - 1 = 6,500$$

$$\text{Fitness C4} = 1/1 + (0,9+0,9+0,9+0,9+0,9+0,1+0,1+0,1) - 1 = 4,900$$

$$\text{Fitness C5} = 1/1 + (0,1+0,9+0,9+0,9+0,9+0,1+0,1+0,9) - 1 = 4,900$$

Dari nilai Fitness diatas, diketahui total fitness adalah 26,900

$$P[C1] = 0,152416357$$

$$P[C2] = 0,241635688$$

$$P[C3] = 0,241635688$$

$$P[C4] = 0,182156134$$

$$P[C5] = 0,182156134$$

3. Seleksi Kromosom induk

Untuk proses seleksi menggunakan Roulette Wheel. Tahap awal menentukan nilai komulatif probabilitas dari setiap kromosom sebagai berikut :

$$C5 = 0,152416357 = 0,152416357$$

$$C5 = 0,152416357 + 0,241635688 = 0,394052045$$

$$C5 = 0,152416357 + 0,241635688 + 0,241635688 = 0,635687732$$

$$C5 = 0,152416357 + 0,241635688 + 0,241635688 + 0,182156134 = 0,817843866$$

$$C5 = 0,152416357 + 0,241635688 + 0,241635688 + 0,182156134 + 0,182156134 = 1$$

Langkah selanjutnya, dengan menggunakan bilangan acak R antara 0 sampai 1. Bilangan acak yang digunakan sebanyak jumlah kromosom, misalkan sebagai berikut :

$$R1 = 0,512$$

$$R2 = 0,401$$

$$R3 = 0,877$$

$$R4 = 0,622$$

$$R5 = 0,817$$

Selanjutnya memilih kromosom ke x sebagai parent dengan syarat $C[x-1] < R < C[x]$. Angka acak $R[1] >$ Nilai Komulatif dari C1, akan tetapi $R[1] <$ dari nilai komulatif C2. sehingga C1 nanti akan dilakukan crossover menjadi :

C1 menjadi
C2:[0.9_0.1_0.9_0.9_0.9_0.1_0.9_0.9_0.9]
C2 menjadi C3
:[0.9_0.9_0.9_0.9_0.9_0.1_0.1_0.9_0.9]
C3 menjadi
C5:[0.1_0.9_0.9_0.9_0.9_0.1_0.1_0.1_0.9]
C4 menjadi
C3:[0.9_0.9_0.9_0.9_0.9_0.1_0.1_0.9_0.9]
C5 menjadi
C4:[0.9_0.9_0.9_0.9_0.9_0.1_0.1_0.1_0.1]

4. Perkawinan Silang (Crossover)
Dalam crossover kita menentukan Probabilitinya (pr), yaitu sebesar 0.5 atau 50%. Hanya kromosom yang nilai R lebih kecil dari 0.5 yang akan bermutasi. Maka Kromosom ke y akan dipilih menjadi induk jika $R[y] < pr$, dari bilangan acak R diatas maka yang dijadikan parent adalah C1 dan C2. Sedangkan C3,C4 dan C5 > 0.5, sehingga tidak dilakukan seleksi. Selanjutnya setelah melakukan pemilihan parent, dilanjutkan menentukan kromosom yang akan di lakukan perkawinan silang dengan pengambilan sejumlah atribut 1-9. Dalam hal ini posisi cut-point (cp) dipilih menggunakan bilangan acak dari 1-9 sesuai banyaknya crossover yang terjadi. Misalnya didapatkan posisi crossover adalah 4, maka kromosom parent akan dipotong mulai gen ke 4 kemudian potongan gen tersebut saling ditukarkan antar parent.

$C1 \times C2$
 $cp(C1) = 4$
 $cp(C2) = 6$
nilai Offspring[1] = $C1 \times C2$ dengan $cp(C1) =$
 $[0.9_0.1_0.9_0.9_0.9_0.1_0.9_0.9_0.9] \times [0.9_0.9_0.9_0.9_0.1_0.1_0.9_0.9]$
 $= [0.9_0.1_0.9_0.9_0.9_0.1_0.9_0.9_0.9]$
nilai Offspring[2] = $C2 \times C1$
dengan $cp(C2) =$
 $[0.9_0.1_0.9_0.9_0.9_0.1_0.9_0.9_0.9] \times [0.9_0.9_0.9_0.9_0.1_0.1_0.9_0.9]$
 $= [0.9_0.9_0.9_0.9_0.9_0.1_0.9_0.9_0.9]$

Populasi baru yang terbentuk dengan hasil crossover yang telah terjadi menjadi :

C1 :[0.9_0.1_0.9_0.9_0.9_0.1_0.9_0.9_0.9]
C2 :[0.9_0.9_0.9_0.9_0.9_0.1_0.9_0.9]
C3 :[0.1_0.9_0.9_0.9_0.9_0.1_0.1_0.1_0.9]
C4 :[0.9_0.9_0.9_0.9_0.9_0.1_0.1_0.9_0.9]
C5 :[0.9_0.9_0.9_0.9_0.9_0.1_0.1_0.1_0.1]

5. Mutasi Kromosom
Jumlah Chromosom yang mengalami mutasi dalam satu populasi ditentukan oleh presentase mutation (pm). Proses mutasi dilakukan dengan cara mengganti satu gen yang terpilih secara acak dengan suatu nilai baru yang didapat secara acak.

Total Gen = (Gen dalam Kromosom)* Jumlah

Kromosom
= 9 gen * 5 kromosom = 80

Selanjutnya dilakukan penentuan posisi gen yang mengalami mutasi dengan cara menggunakan bilangan acak anantara 1 sampai 45 (sesuai total gen). Lalu menentukan presentase mutation, misalnya sebesar 10%. Maka jumlah gen yang mengalami mutasi 10% dari 45 adalah 4-5 gen. Kemudian menggunakan bilangan acak 0.00 – 1.00 untuk mengganti 4 gen yang dipilih secara acak juga. Misalnya posisi gen 10-14 akan diganti dengan sejumlah nilai secara berurutan [0.1_0.9_0.9_0.9_0.1]. Dengan demikian kromosom ke 2 pada gen ke 2 sampai kromosom ke 2 pada gen ke 10. maka hasil yang didapat adalah :

C1 :[0.9_0.1_0.9_0.9_0.9_0.1_0.9_0.9_0.9]
C2 :[0.1_0.9_0.9_0.1_0.9_0.9_0.1_0.9_0.9]
C3 :[0.1_0.9_0.9_0.9_0.9_0.1_0.1_0.1_0.9]
C4 :[0.9_0.9_0.9_0.9_0.9_0.1_0.1_0.9_0.9]
C5 :[0.9_0.9_0.9_0.9_0.9_0.1_0.1_0.1_0.1]

Selanjutnya kromosom ke 3 di uji agar menjadi kromosom yang diinginkan. Jika kromosom ke 3 tidak memenuhi syarat, maka semua kromosom dalam populasi ini akan mengalami proses yang sama seperti generasi sebelumnya yaitu proses evaluasi, seleksi, crossover dan mutasi yang kemudian akan menghasilkan kromosom-kromosom baru untuk generasi selanjutnya. Gen yang sama dianggap tidak digunakan pada metode selanjutnya dan dibuat bobotnya menjadi 0. Dan Gen yang berbeda pada setiap kromosom digunakan pada metode nantinya dan dibuat bobot menjadi 1.

attribute	weight
NAMA SISWAI	0
STATUS ORANG TUA	0
PEKERJAAN & PENGHASILAN ORANG TUA	0
Rumah, kontrakan, tanah	1
peralatan rumah	1
kendaraan, jumlah	1
tabungan keluarga	1
perhiasaan orang tua	0
handphone orang tua	0
UANG JAJAN	0
Class Menerima beasiswa	1

Gambar 9. Bobot dalam Algoritma Genetika

Setelah dilakukan pembobotan maka terdapat beberapa atribut yang tidak digunakan, setelah itu diseleksi dengan menggunakan uji kelayakan dengan taraf kesalahan sebesar 1% (Sugiyono,2007) yang menghasilkan sample sebanyak 154 dari 200 siswa yang mendapatkan beasiswa.

	true Y	true N	class precision
pred Y	25	7	78.12%
pred N	3	25	89.29%
class recall	89.29%	78.12%	

Gambar 10. Hasil Akurasi kedua algoritma

Pada gambar ini, akurasi yang dihasilkan dari uji 154 siswa adalah 83,33%. Adapun perhitungan manual seperti berikut :

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{25 + 25}{25 + 25 + 7 + 3} = 0,83 = 83,33\%$$

$$\text{Recall / TP Rate} = \frac{TP}{TP + FN} = \frac{25}{25 + 3} = 0,89 = 89,29\%$$

$$\text{Recall / TN Rate} = \frac{TN}{TN + FP} = \frac{25}{25 + 7} = 0,78 = 78,12\%$$

$$\text{NPV} = \frac{TN}{TN + FN} = \frac{25}{25 + 3} = 0,89 = 89,29\%$$

$$\text{PPV} = \frac{TP}{TP + FP} = \frac{25}{25 + 3} = 0,78 = 78,12\%$$

KESIMPULAN

Penelitian ini dilakukan agar mendapatkan akurasi tertinggi dari kedua algoritma yang digunakan untuk memprediksi siswa apakah mendapatkan beasiswa pendidikan atau tidak pada SMP Utama Depok. Berdasarkan algoritma naïve bayes didapatkan akurasi sebesar 77,50% dalam memprediksi penerima beasiswa berdasarkan kriteria – kriteria yang terdapat pada peserta didik. Untuk optimasi digunakan algoritma genetika, nilai akurasi mencapai 83,33%. Hal ini membuktikan penggunaan dua algoritma, yaitu algoritma naïve bayes dan algoritma genetika memberikan hasil akurasi tinggi. Penelitian ini masih jauh dari kata sempurna karenanya di sarankan sebagai berikut:

1. Melakukan pengembangan penelitian dengan algoritma yang berbeda sehingga mendapatkan nilai akurasi yang lebih optimal seperti menggunakan algoritma optimasi Particle Swarm Optimization dan yang lainnya.
2. Perlu adanya penetapan dan penambahan kriteria penerima beasiswa agar lebih tepat sasaran.

REFERENSI

- Kusrini & Emha Taufiq Luthfi. (2009). *Algoritma Data Mining (1st Publis)*. Yogyakarta: Andi Offset.
- Meilani, B. D., & Susanti, N. (2015). Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Siswa Dengan Metode Naive Bayes. *Jurnal Ilmiah NERO*, 1(3), 182–189.

- Murniasih, E. (2008). *Winning A Scholarship*. Jakarta: Gagas Media.
- Murniasih, E. (2009). *Buku Pintar Beasiswa*. Jakarta: Gagas Media.
- Obitko, M. (1998). *Genetic Algorithms*. Retrieved from <https://www.obitko.com/tutorials/genetic-algorithms/index.php>
- Prabowo Pudjo Widodo, Rahmadya Trias Handayanto, H. (2013). *Penerapan Data Mining Dengan Matlab (1st ed.)*. Bandung: REKAYASA SAINS Bandung.
- Prasetyo, E. (2012). *Data Mining: Konsep dan Aplikasi menggunakan MATLAB (1st ed)*. Yogyakarta: Andi.
- Prasetyo, E. (2014). *Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab*. Yogyakarta: Andi Offset.
- Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Tullah, R., Mariana, A. R., & Baskoro, D. (2018). *Sistem Pendukung Keputusan Pemilihan Calon Penerima Beasiswa Bidikmisi Menggunakan Metode AHP Dan TOPSIS Pada STMIK Bina Sarana Global*, 8(2).

PROFIL PENULIS

Nining Suryani. Lahir di Jakarta pada tanggal 6 Januari. Bergabung dengan AMIK BSI sejak tahun 2002 memperoleh gelar Sarjana Komputer STMIK KUWERA tahun 2003 dan memperoleh gelar Magister Komputer (M.Kom) Program Pasca Sarjana STMIK Nusa Mandiri tahun 2010, menggeluti bidang ilmu komputer. Memiliki penulisan dalam bentuk jurnal ilmiah diantaranya *Emoticon Untuk Pengukuran Kepuasan Kerja Karyawan Pada Perusahaan*, *Paradigma* Maret 2011, *Maturity Level Dan Management Awareness Pada Perancangan Tata Kelola Teknologi Informasi Perusahaan Tunas Kreasi Mandiri*, *Pilar* Maret 2013, *Implementasi Sistem Informasi Pembelajaran Berbasis e-learning Studi Kasus (SMK Nuurul Bayan Kalapanunggal Sukabumi)* *IBI* September 2013, *Paradigma* Maret 2014, *Pembelajaran Interaktif Bahasa Inggris Bagi Siswa Kelas 2 Sekolah Dasar (Studi Kasus Mi Nurul Falah Ciater)*, *Paradigma* September 2014

Evy Priyanti, M.Kom Lahir di Jakarta pada tahun 1986 dan menyelesaikan pendidikan Diploma III Pada jurusan Manajemen Informatika dan Komputer di Bina Sarana Informatika Jakarta pada tahun 2007, pada tahun 2008 menamatkan Studi Strata satu pada STMK Kuwera dan Strata Dua pada bidang ilmu komputer di STMIK Nusa Mandiri. Menjadi Asisten Komputer pada AMIK BSI Jakarta pada tahun 2005 dan diangkat menjadi instruktur pada tahun 2008. Tahun 2017 berhasil naik jabatan menjadi asisten ahli dan staff akademik. Saat ini menjadi anggota konsorsium sistem informasi dan juga anggota Asosiasi Dosen Indonesia.