KILAT
Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169

Analisis Kelayakan Lokasi Promosi Dalam Penerimaan Mahasiswa Baru (PMB) Dengan Algoritma Naïve Bayes & Decission Tree C4.5

Sekar Wulandari[1]; Mohammad Iqbal[2]

[1, 2]MMSI, Universitas Gunadarma

[1] sekarwulandarie@gmail.com
[2] iqbalgoh@gmail.com

ABSTRACT

Competition for new student admissions in every public and private tertiary institution is currently growing rapidly every year, some spend a lot of money on promotional activities, to assist institutions / institutions in obtaining recommendations for the feasibility of promotion locations based on several measurement criteria using the classification algorithms contained in data mining . The algorithm used to compare the measurement of the feasibility of the promotion location of the city and district of Bekasi is Naïve Bayes and Decission Tree C4.5 using four parameters including the number of students in one sub-district, the number of students in one sub-district, the distance of location and last year's enthusiasts using 35 regions / sub-districts in Bekasi city and district. measurement results using the rapidminner, the accuracy value of the Naïve Bayes algorithm is 91.43% and the Decission Tree C4.5 is 94.29%.

Keywords: *Classification, Naïve Bayes, C4.5*

ABSTRAK

Persaingan penerimaan mahasiswa baru disetiap perguruan tinggi negeri maupun swasta saat ini semakin pesat setiap tahunnya, ada yang menghabiskan banyak biaya untuk kegiatan promosi, untuk membantu institusi/lembaga dalam memperoleh rekomendasi kelayakan lokasi promosi berdasarkan dari beberapa pengukuran kriteria menggunakan algoritma classification yang terdapat dalam data mining. Algoritma yang digunakan untuk membandingkan pengukuran kelayakan lokasi promosi wilayah kota dan kabupaten bekasi adalah Naïve Bayes dan Decission Tree C4.5 dengan menggunakan empat parameter meliputi Jumlah peserta didik dalam satu kecamatan, jumlah siswa dalam satu kecamatan, jarak lokasi dan peminat tahun lalu dengan menggunakan 35 wilayah/kecamatan yang ada di kota dan kabupaten Bekasi. Dari hasil pengukuran menggunakan rapidminner diperoleh nilai akurasi pada algoritma Naïve Bayes sebesar 91,43% dan Decission Tree C4.5 sebesar 94,29%.

Kata kunci: *Klasifikasi, Naïve Bayes, C4.5*

Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169

1. PENDAHULUAN

Pendidikan di Indonesia saat ini termasuk dalam program nawacita yang dicanangkan oleh Presiden Republik Indonesia, berdasarkan data dari Badan Pusat Statistik (BPS) [1] Angka Partisipasi Kasar (APK) Perguruan Tinggi (PT) terjadi peningkatan prosentase sejak tahun 2015 sampai dengan tahun 2019 secara nasional, pada tahun 2019 tercatat sebanyak 30, 28 % lulusan SMA sederajat yang melanjutkan pendidikan ke jenjang perguruan tinggi secara nasional dan 25,15 % lulusan SMA sederajat yang melanjutkan pendidikan ke jenjang perguruan tinggi untuk wilayah provinsi Jawa Barat. berdasarkan data terdahulu bahwa angka partisipasi lulusan kelas XII di Kota [2] dan Kabupaten Bekasi [3] untuk melanjutkan pendidikan lanjutan keperguruan tinggi juga tidak lebih dari 23,48 % pada tahun 2015.

Persaingan penerimaan mahasiswa baru disetiap perguruan tinggi negeri maupun swasta saat ini semakin pesat setiap tahun nya, ada yang menghabiskan banyak biaya untuk kegiatan promosi dengan meggunakan berbagai program seperti murahnya biaya perkuliahan, adanya program sertifikasi internasional, kemudahan mendapatkan pekerjaan setelah lulus dan lain sebagainya.

Sekolah Tinggi Manajemen dan Informatika (STMIK) Bani Saleh melakukan kegiatan promosi penerimaan mahasiswa baru (PMB) dimana prioritas utama dari promosi tersebut adalah untuk mendapatkan calon mahasiswa sesuai dengan daya tampung yang telah disediakan. Untuk mendapatkan calon mahasiswa pada tahun akademik 2020/ 2021 sesuai dengan yang diharapkan, tentunya pihak lembaga sebagaimana tahun tahun sebelumnya akan melakukan promosi ke berbagai tempat, baik di wilayah Bekasi ataupun daerah sekitar Bekasi. Usaha yang dilakukan untuk penerimaan tahun akademik 2020/ 2021 bagaimana promosi dapat dilakukan secara terarah dan tepat sasaran agar semua upaya dan pembiayaan yang dikeluarkan untuk promosi menjadi efektif dengan memberikan hasi penerimaan mahasiswa baru yang meningkat.

Dengan mengumpulkan data yang bersumber dari internal dan eksternal sebagai parameter pengukuran kelayakan, penerapan data mining dapat digunakan dengan mengolah data. Dari sumber eksternal pengambilan kriteria jarak lokasi promosi, jumlah sekolah dalam suatu kecamatan, banyak siswa dalam suatu kecamatan yang diperoleh dari situs website resmi yang diunduh pada situs resmi kemendikbud.go.id [4] sementara peminat tahun lalu pada STMIK Bani Saleh mengambil data PMB empat tahun terakhir yaitu TA. 2016 sampai dengan tahun 2019. Terjadinya perubahan minat mahasiswa dalam memilih waktu kuliah yang semula didominasi mahasiswa shift/ perkuliahan malam menjadi perkuliahan pagi dimana rata-rata perkuliahan pagi adalah mahasiswa yang tidak/belum bekerja sementara rata-rata mahasiswa yang memilih waktu perkuliahan di malam hari adalah pekerja, sehingga perlu diadakan analisis untuk perubahan metode promosi.

2. METODE PENELITIAN

2.1. Pengertian Data Mining

Data mining merupakan proses menemukan pola yang akan memberikan wawasan serta model deskriptif, dapat dipahami dan memberikan prediksi data data yang berskala besar [15]. adapun tahapan dalam data mining antara lain adalah:
beberapa Tahapan dalam data mining antara lain [14]:

1. Pembersihan data

   Data-data yang tidak relevan lebih baik dibuang karena keberadaannya bisa mengurangi mutu atau akurasi dari hasil *data mining* nantinya.
2. Integrasi data

Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169

Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya.

3. Tranformasi data
   Transformasi dan pemilihan data ini juga menentukan kualitas dari hasil data mining nantinya karena ada beberapa karakteristik dari teknik-teknik data mining tertentu yang tergantung pada tahapan ini.
4. Aplikasi Teknik data mining
   Aplikasi teknik *data mining* sendiri hanya merupakan salah satu bagian dari proses *data mining*.
5. Evaluasi Pola yang ditemukan
   Dalam tahap ini hasil dari teknik *data mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti : menjadikannya umpan balik untuk memperbaiki proses *data mining*, mencoba teknik *data mining* lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

A. Teknik Data Mining

Ada banyak jenis teknik analisa yang dapat digolongkan dalam *Data mining* baik yang menggunakan metode deskriptif ataupun yang menggunakan metode prediktif [6] salah satunya adalah *Classification*, merupakan metode prediktif dimana teknik ini bekerja dengan menemukan model atau fungsi dari sekumpulan data yang ada. Adapun yang termasuk dalam algoritma classification antara lain adalah :

1. Algoritma Naïve Bayes
   Adalah metode memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya. NBC merupakan pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Teori keputusan Bayes merupakan pendekatan statistik yang fundamental dalam pengenalan pola (pattern recognition) [5]. Teori Bayesian pada dasarnya adalah kemungkinan kejadian di masa depan yang bisa dihitung dengan menentukan frekuensi pengalaman sebelumnya. Penggunaan algoritma Bayes dalam hal klasifikasi harus mempunyai masalah yang bisa dilihat statistiknya [6]. Untuk menghitung nilai kelas yang akan dibandingkan (ya atau tidak), dilakukan perhitungan probabilitas P(Vj):

$$P(Vj) = \frac{|docj|}{|contoh|}$$

   Dimana docj adalah banyaknya dokumen yang memiliki kategori j dalam pelatihan, sedangkan Contoh banyaknya dokumen dalam contoh yang digunakan untuk pelatihan. Untuk nilai P(Wk|Vj), yaitu probabilitas kata Wk dalam kategori j di tentukan dengan :

$$P(Wk|Vj) = \frac{Nk+1}{N+|vocabulary|}$$

Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169

Dimana nk adalah frekuensi munculnya kata wk dalam dokumen yang berkategori vj ditambah 1, hal ini berfungsi untuk menghindari angka 0 dalam data atau di sebut Laplace Smoothing, sedangkan nilai n adalah banyaknya seluruh kata dalam dokumen berkategori vj, dan vocabulary adalah banyaknya kata dalam contoh pelatihan.

2. Algoritma C4.5

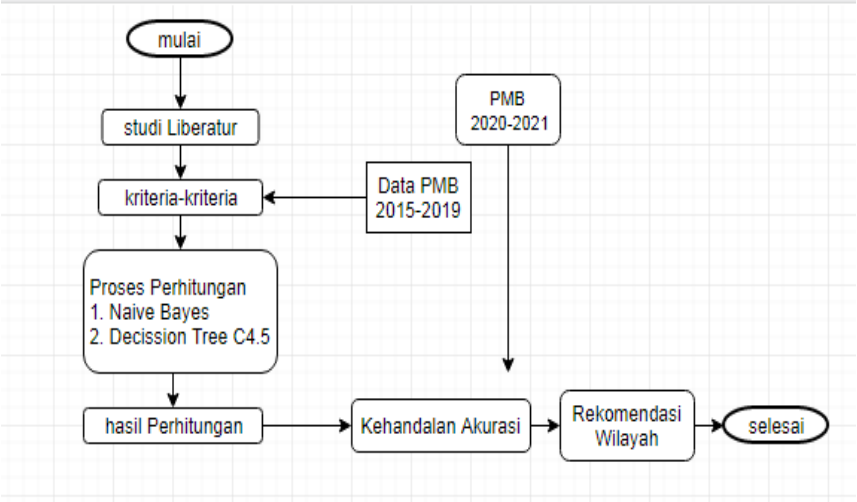
   Merupakan salah satu algoritma yang melakukan klasifikasi atau pengelompokan pada data set.dasar dari algoritma C4.5 adalah pembentukan pohon keputusan (Decission Tree). Cabang cabang dari pohon keputusan merupakan pernyataan klasifikasi sedangkan daun daunnya merupakan kelas kelas atau kelompoknya [6]. Secara umum konsep algoritma C4.5 adalah sebagai berikut :

   1. Pilih variable sebagai akar
   2. Buat cabang untuk masing masing nilai
   3. Bagi kasus dalam cabang
   4. Ulangi proses untuk masing masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

2.2. Penelitian sebelumnya

Penelitian yang dilakukan oleh Maryamah., Moh.Fadel Asikin Cs. Pada tahun 2018 [10] menghasilkan klasifikasi yang akurat, sehingga dapat mempermudah seorang dosen dalam memilih asisten dengan kualitas yang baik, Ilka Zufria, pada tahun 2018 [13] menghasilkan rekomendasi untuk menentukan potensi lokasi promosi dan pemasaran. Penelitian yang dilakukan oleh Oktariani Nurul Pratiwi pada tahun 2016 [11] hasil uji coba didapatkan Naïve Bayes mampu mengelompokan siswa lebih baik dengan nilai akurasi 70,37 %. Penelitian yang dilakukan oleh yulia fahmi pada tahun 2019 [8] diketahui bahwa nilai kepercayaan merupakan salah satu faktor yang penting untuk membuka usaha pada suatu lingkungan. Pada tahun 2018, dilakukan penelitian oleh Haditzah annur [9] diperoleh hasil dengan tingkat akurasi sebesar 73% atau termasuk dalam kategori Good. Sementara nilai Precision sebesar 92% dan Recall sebesar 86%. Penelitian dilakukan oleh Sri wahyuningsih cs, pada 2018 [12] menghasilkan Algoritma Decission Tree (J48) dengan nilai akurasi sebesar 92,21 %, Algoritma K-Nearest Neighbor memiliki tingkat akurasi sebesar 81,82 % dan Algoritma Naïve Bayes memiliki tingkat akurasi sebesar 81,83 %.

2.3. Tahapan penelitian



Gambar 1. Diagram Alir Penelitian

Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169

Berikut penjelasan dari diagram alir tahapan penelitian :
Adapun tahapan yang dilakukan dalam penelitian ini antara lain :

1. Studi Literatur

   Kegiatan promosi penerimaan mahasiswa baru merupakan agenda rutin yang selalu dilakukan oleh sebuah institusi sebagai bentuk keberlanjutan dari kegiatan akademik. Perlunya melakukan kegiatan promosi agar dapat dilakukan secara terarah dan tepat sasaran supaya semua upaya dan pembiayaan yang dikeluarkan untuk promosi menjadi efektif dengan memberikan hasil penerimaan mahasiswa baru yang meningkat. Berdasarkan pemikiran tersebut, maka diperlukan upaya dan langkah strategis agar penerimaan mahasiswa baru sesuai dengan daya tampung. Sebagai langkah awal maka dilakukan studi literatur mengenai kelayakan lokasi promosi untuk lebih memahaminya. Studi literatur dilakukan dengan membaca serta mempelajari jurnal-jurnal ilmiah, paper, artikel, buku, serta sumber ilmiah lainnya.

2. Data dan Kriteria

   Dalam melakukan proses penelitian tentu membutuhkan data yang cukup banyak dan sesuai dengan yang dibutuhkan, dalam penelitian ini peneliti menggunakan Data sekunder dari berbagai sumber antara lain:

   1. Kriteria Jumlah sekolah dan Jumlah Peserta didik Kota dan Kabupaten di Bekasi berasal dari situs resmi kemendikbud.go.id Pengambilan data pada tahun 2019
   2. Kriteria Jarak berasal dari Google Map
   3. Kriteria Peminat tahun lalu berasal dari Puskom STMIK Banisaleh, data penerimaan mahasiswa baru STMIK Bani Saleh Tahun angkatan 2015 sampai dengan 2019 sebagai data training dan penerimaan mahasiswa tahun angkatan 2020 sebagai data testing.

3. Tranformasi Data

   Proses data mining Knowledge Discovery in database (KDD)
   Proses dalam tahapan Knowledge Discovery in database prepocesssing sebagai berikut :

4. Data selection

   Pada tahap ini data akan diseleksi dengan cara melihat kecenderungan data/kesesuaian data topik/ judul penelitian yang akan diteliti, adapun data yang akan diteliti diperoleh dari situs web kemendikbud.go.id untuk data jumlah peserta didik dalam satu kecamatan dan jumlah sekolah dalam satu kecamatan, google map untuk data jarak dan data PMB untuk data peminat tahun lalu.

   a. Data selection jumlah peserta didik

      Data Jumlah peserta didik diambil melalui situs resmi pada yaitu kemendikbud.go.id [4] yang terdiri dari sekolah menengah atas untuk SMA dan SMK baik negeri maupun swasta yang berada di wilayah Kota sebanyak 12 kecamatan dan Kabupaten Bekasi sebanyak 23 kecamatan. Sehingga total wilayah adalah 35 kecamatan.

2.3. Aplikasi Pengolahan

Aplikasi Pengolahan dalam pengukuran kelayakan lokasi promosi adalah ***Rapidminner 5.3***

3. HASIL DAN PEMBAHASAN

3.1. Pembersihan Data

a. Tahap *Pembersihan data untuk kriteria Peserta didik kota dan Kabupaten*

Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169

Dari data yang diakses melalui kemendikbud.go.id [4] jumlah sekolah dalam satu kecamatan Data Jumlah sekolah diambil melalui situs resmi kementerian pendidikan dan kebudayaan yaitu yang terdiri dari sekolah menengah atas untuk SMA dan SMK baik negeri maupun swasta yang berada di wilayah Kota sebanyak 12 kecamatan dan Kabupaten Bekasi sebanyak 23 kecamatan. Sehingga total wilayah adalah 35 kecamatan. Adapun data yang diambil adalah jumlah peserta didik seperti yang tercantum dalam tabel 1.8

Tabel 1. Data Jumlah peserta didik di Kota dan Kabupaten Bekasi

No	Wilayah	SMA			SMK			TOTAL
		Jml	L	P	Jml	L	P	
1	Kec. Tambun Selatan	8199	3.463	4736	3635	2.278	1.357	11834
2	Kec. Cikarang Utara	2278	903	1375	1146	725	421	3424
3	Kec. Cikarang Selatan	2242	757	1.485	2934	1.75	1.184	5176
4	Kec. Cikarang Barat	424	152	272	1827	652	1.175	2251
5	Kec. Babelan	474	202	272	2929	1.626	1.303	3403

b. Tahapan Integrasi data

Pada tahap integrasi data merupakan tahap penggabungan data wilayah, Jumlah Peserta didik, Jumlah sekolah, Jarak dan Peminat tahun lalu , data integrasi adalah data yang berasal dari berbagai sumber kemudian digabungkan menjadi dataset. Penggabungan data dalam integrasi ini terdiri dari empat parameter seperti pada tabel 2

Tabel 2. Data Integration

No	Wilayah	Total Peserta Didik	Total Sekolah	Jarak	Peminat tahun lalu
1	Kec. Rawalumbu	6130	32	3.6 km	13
2	Kec. Bekasi Timur	3299	34	3 km	116
3	Kec. Bekasi Utara	6090	32	14.6 km	78
4	Kec. Bekasi Barat	1169	19	9 km	33
5	Kec. Bekasi Selatan	6134	18	6.9 km	65
6	Kec. Pondokgede	3314	26	10.7 km	7
7	Kec. Jatiasih	4662	23	7.5 km	11
8	Kec. Mustika Jaya	5586	22	7 km	58
9	Kec. Medan Satria	3237	19	12.7 km	20
10	Kec. Bantargebang	3729	13	5.5 km	24

Data integrasi meliputi empat kriteria yaitu wilayah untuk kota dan kabupaten yang berjumlah 35 kecamatan, peserta didil dan total sekolah serta jarak, sedangkan peminat tahun lalu berjumlah 900 data yang terdiri dari 4 tahun lalu dan 180 data yang terdiri dari data tahun terakhir PMB 2020.

3.2. Data Tranformasi

Adapun dalam tahap ini data akan diubah menjadi bentuk yang sesuai untuk proses data mining.

Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169

Tabel 3. Nilai bobot

Atribut	Keterangan	Nilai
Jumlah Sekolah	> 20 Sekolah	3
	11 – 20 Sekolah	2
	0 – 10 Sekolah	1
Jumlah Siswa	> 5000 Siswa	3
	2001 – 5000 Siswa	2
	<= 2000	1
Jarak	< 7 Km	3
	7 - 20 Km	2
	> 20 Km	1
Peminat Tahun Lalu	> 20 Sekolah	3
	11 – 20 Sekolah	2
	0 – 10 Sekolah	1

Tabel 4. Nilai label

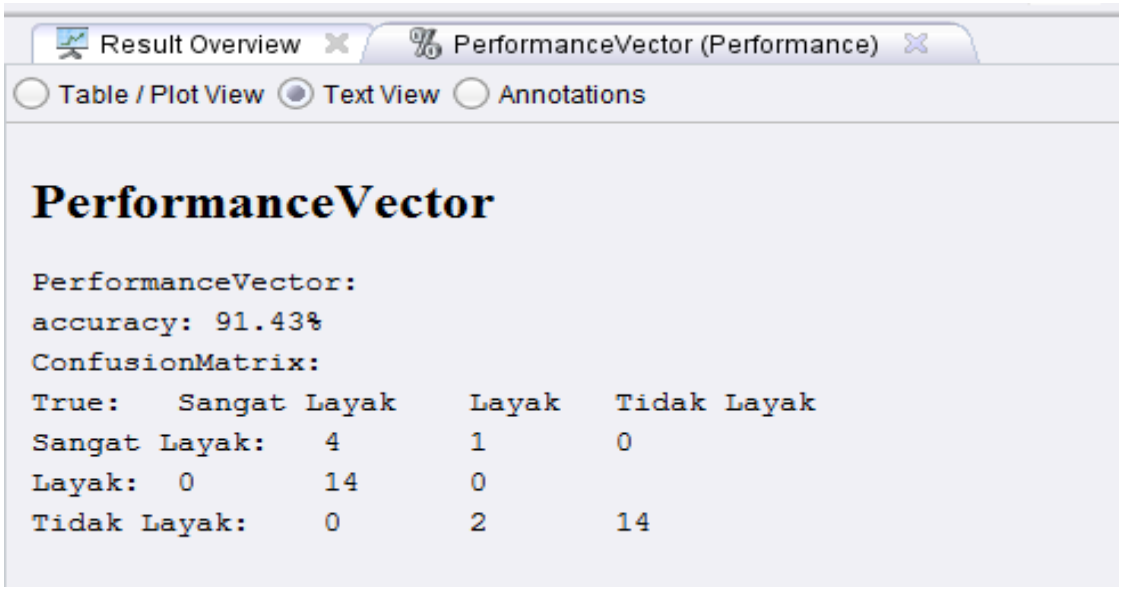
Label	Total Nilai	Nama Label
Status Kelayakan	>=10	Sangat Layak
	<7 x <10	Layak
	< 7	Tidak Layak

Tabel 5. Tabel Tranformation

No	Wilayah	Jumlah Peserta Didik	Jumlah Sekolah	Jarak	Peminat Tahun Lalu	Status
1	Kec. Rawalumbu	3	3	3	2	Sangat Layak
2	Kec. Bekasi Timur	2	3	3	3	Sangat Layak
3	Kec. Bekasi Utara	3	3	2	3	Sangat Layak
4	Kec. Bekasi Barat	1	2	2	3	Layak
5	Kec. Bekasi Selatan	3	2	2	3	Layak
6	Kec. Pondokgede	2	3	2	1	Layak
7	Kec. Jatiasih	2	3	2	2	Layak
8	Kec. Mustika Jaya	3	3	2	3	Sangat Layak
9	Kec. Medan Satria	2	2	2	2	Layak
10	Kec. Bantargebang	2	2	3	2	Layak

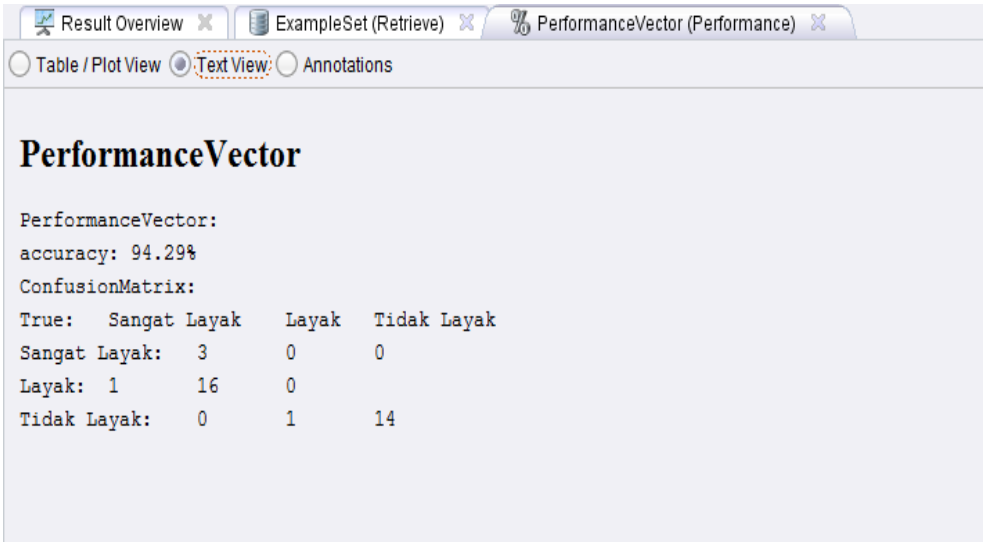
Perhitungan dengan menggunakan Rapid Minner 5.3 diperoleh hasil dari Pervormance vector berupa text view menghasilkan akurasi pada Naïve Bayes menghasilkan angka sebesar 91,43 % dengan status sangat layak berjumlah 4 dan layak berjumlah 14 kecamatan.

Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169



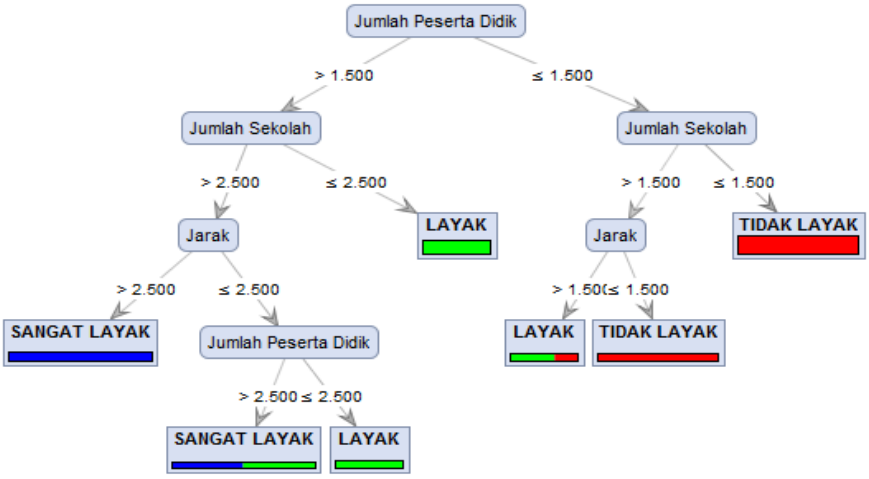
Gambar 2. Pervormance Vector Naïve Bayes

Hasil dari Pervormance vector berupa text view menghasilkan accuracy pada C 4.5 menghasilkan angka sebesar 91,43 % dengan status sangat layak adalah 3 dan layak berjumlah 16 Kecamatan.



Gambar 3 Pervormance Vector C 4.5

Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169



Gambar 4. Pohon Keputusan C 4.5

4. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Dari hasil penelitian yang dilakukan mulai dari tahap awal hingga proses pengujian aplikasi dapat disimpulkan bahwa :

a. Algoritma yang digunakan untuk membandingkan pengukuran kelayakan lokasi promosi wilayah kota dan kabupaten bekasi adalah naïve bayes, Decission Tree C4.5 dengan menggunakan empat parameter meliputi Jumlah peserta didik, jumlah siswa, jarak lokasi dan peminat tahun lalu dengan menggunakan 35 wilayah yang ada di kota dan kabupaten Bekasi dengan menggunakan 900 sampel data training dan 180 data testing dan Implementasi pengolahan data mining dengan menggunakan rapidminner 5.300
b. Dari perhitungan menggunakan aplikasi rapidminner bahwa nilai accuracy dari C4.5 lebih baik dibandingkan Naïve Bayes dimana diperoleh nilai akurasi sebesar 94,29% dan Naïve Bayes sebesar 91,43%
c. Dari hasil pengukuran kedua model algoritma mempunyai kehandalan yangg cukup, dari yg sangat layak dan layak berdasarkan 4 tahun lalu, ketika di validasi dengan data test (100 data) dari data PMB Tahun Akademik 2020 ternya hasilnya masih tetap layak dan sangat layak.

4.2. Saran

Berdasarkan hasil penelitian perbandingan pengukuran lokasi promosi dengan menggunakan dua algoritma dalam classification dapat memberikan rekomendasi kelayakan lokasi promosi yang dilakukan tim promosi Penerimaan Mahasiswa Baru agar lebih terarah agar upaya dan pembiayaan untuk promosi menjadi efektif dengan memberikan hasil penerimaan mahasiswa baru yang meningkat, namun ada beberapa hal yang perlu disarankan Pengukuran kelayakan lokasi promosi dapat dilakukan dengan algoritma yang berbeda dalam data mining.

Vol. 10, No. 1, April 2021, P-ISSN 2089-1245, E-ISSN 2655-4925
DOI: https://doi.org/10.33322/kilat.v10i1.1169

DAFTAR PUSTAKA

[1] Badan Pusat Statistik 2020. https://www.bps.go.id/ (12 Juni 2020 09.00)
[2] Badan Pusat Statistik Kota Bekasi 2020 https://bekasikota.bps.go.id/ (12 Juni 2020 10.00)
[3] Badan Pusat Statistik Kab Bekasi 2020 https://bekasikab.bps.go.id/ (12 Juni 2020 12.00)
[4] kemendikbud 2020. https://dapo.dikdasmen.kemdikbud.go.id/sp/2/026500 (12 Juni 2020 08.30)
[5] Anggarwal C C (2015) Data Mining. New York: Springer.
[6] J. Zaki, Mohammed. 2014. Data mining and analysis. New York: Cambridge University Press.
[7] Chapman & Hall (2015) Data Classification Algorithm & Application. U.S.A: CRC Press.
[8] Yulia Fahmi (2019) Prediksi Pembukaan Peluang Usaha di Kecamatan Medan Amplas menggunakan Algoritma Naïve Bayes. Riau Jurnal Of Computer science Vol 05 No. 01.
[9] Annur, Haditzah (2018) Klasifikasi Masyarakat miskin menggunakan metode naïve bayes. ILKOM Jurnal Ilmiah Volume 10 Nomer 2 agustus 2018.
[10] Maryamah., Moh.Fadel Asikin Cs (2018) Implementasi metode Naïve Bayes Classifier Untuk Seleksi Asisten Praktikum Pada Simulasi Hadoop Multinde Cluster.Jurnal Teknologi Informasi dan Ilmu Komputer.
[11] Nurul O, (2016) Analisa Perbandinga Algoritma K-Means, Decision Tree, dan Naïve Bayes Untuk Sistem Pengelompokan Siswa Otomatis. Jurnal Ilmiah Teknologi Informasi Terapan.
[12] Wahyuningsih, Sri.cs (2018) Perbandingan Metode K-Nearest Neighbor, Naïve Bayes dan Decission Tree untuk memprediksi kelayakan pemberian kredit. Konferensi Nasinal Sistem Informasi.
[13] Zufria I, (2018) Penentuan Lokasi Promosi Calon Mahasiswa Baru Perguruan Tinggi Swasta Berbasis AHP (Analytical Hierarchy Process) journal Of Islamic Science and Technology.
[14] F. Gorunescu 2011. Data Mining concept model Technique Craiova, Romania: Springer.
[15] Jiawei Han, Micheline kamber, 2012. Data Mining concept and techniques Third Edition, USA: Elseiver inc.