Volume 11 Number 1 May 2021 Page. 67-74
Journal Homepage : http://teknois.stikombinaniaga.ac.id/index.php/JBS
DOI Link : http://doi.org/10.36350/jbs.v11i1

ISSN : 2087 - 3891 (PRINT)
ISSN : 2597 - 8918 (DIGITAL)
Jurnal Ilmiah Teknologi - Informasi & Sains
TEKNOIS
Sekolah Tinggi Ilmu Komputer Binaniaga
Change, Continuous Quality Improvement, De'bottlenecking, Enhancement... Do It!

Article DOI : 10.36350/jbs.v11i1.104
Received: March; Accepted : April; Published : May

Penerapan Metode Naïve Bayes untuk Prediksi Minat Baca Berdasarkan Usia

Lis Utari[1*], Yawmil Ulfah[2]
[1]Sistem Informasi/STIKOM Binaniga
Email: lis_utari@yahoo.com
[2]Sistem Informasi/STIKOM Binaniga
Email: yawmilulfah8@gmail.com

ABSTRACT

Reading is one of the activities favored by all circles of society, and a place that can be visited for reading activities is the library. One of them is the Pajajaran Sports Hall Library, based on the data obtained from more than 100 visitors who come to visit the Pajajaran Sports Hall Library each year. However, visitors who borrow do not necessarily have a high reading interest. In this study, an application was made that can predict library visitors who have a high reading interest. In order to predict future reading interest using the method Naive Bayes. It uses variables obtained from borrowing data and carried out preprocessing and obtained several variables used to make predictions, namely name, book category, number of late days, age group, and level of reading interest. This is done to predict the reading interest of visitors to the Pajajaran GOR Library, who have a high level of reading interest and a low level of reading interest. So that the librarian can find out the level of reading interest of the visitors, and the collection of books in the Pajajaran GOR Library can be increased based on the preferred book category of all ages. The accuracy test has also been carried out using the formula confusion matrix with an accuracy value of 80%, which means that the application that has been made is suitable for use.

Keywords*: Naive Bayes Method; Library; Prediction; Reading Interest.*

ABSTRAK

Membaca merupakan salah satu kegiatan yang digemari oleh semua kalangan masyarakat, dan tempat yang dapat dikunjungi untuk kegiatan membaca adalah Perpustakaan. Salah satunya adalah Perpustakaan GOR Pajajaran, berdasarkan dari data pengunjung yang didapat mencapai 100 lebih yang datang berkunjung ke Perpustakaan GOR Pajajaran untuk setiap tahunnya. Tetapi pengunjung yang melakukan peminjaman belum tentu memiliki minat baca yang tinggi. Pada penelitian ini, dibuat sebuah aplikasi yang dapat memprediksi pengunjung perpustakaan yang memiliki minat baca tinggi. Agar dapat dilakukan prediksi minat baca kedepannya dengan menggunakan metode Naive Bayes. Di dalamnya menggunakan variabel-variabel yang didapat dari data peminjaman dan dilakukan preprocessing dan didapat beberapa variabel yang digunakan untuk melakukan prediksi yaitu nama kategori buku, jumlah hari telat, kelompok usia, dan tingkat minat baca. Hal ini dilakukan untuk memprediksi minat baca pengunjung Perpustakaan GOR Pajajaran, yang memilliki tingkat minat baca yang tinggi dan tingkat minat baca yang rendah. Agar bagian petugas perpustakaan dapat mengetahui tingkat minat baca para pengunjung, dan dapat ditingkatkan koleksi buku di Perpustakaan GOR Pajajaran berdasarkan kategori buku yang disukai dari setiap kalangan usia. Telah dilakukan uji akurasi dengan menggunakan rumus confution matrix dengan hasil nilai akurasi sebesar 80%, yang berarti aplikasi yang telah dibuat layak untuk digunakan.

Keywords*: Metode Naive Bayes; Perpustakaan; Prediksi; Minat Baca.*

Volume 11 Number 1 May 2021 Page. 67-74
Journal Homepage : http://teknois.stikombinaniaga.ac.id/index.php/JBS
DOI Link : http://doi.org/10.36350/jbs.v11i1

A. PENDAHULUAN

1. Latar Belakang

Saat ini, perpustakaan termasuk salah satu tempat yang mengatur serta menyimpan banyak koleksi bahan pustaka sedemikian rupa untuk sebagai sumber informasi yang dapat digunakan. Selain itu perpustakaan memiliki banyak kategori buku diantaranya ada kategori buku umum, filsafat, psikolog, agama, sosial, bahasa, teknologi, seni, rekreasi dan yang lainnya. Menurut C. Larasati Milburga (986), menyatakan bahwa perpustakan dapat dijadikan sebagai sarana untuk belajar mengajar sehingga dapat memperoleh informasi dan pengetahuan didalamnya**.**

Sandjaja (2005:3), menyatakan bahwa minat baca yaitu seseorang yang memiliki ketertarikan atau kesukaan terhadap kegiatan membaca. Minat baca dapat dikatakan adanya keinginan dan ketertarikan yang dilakukan secara terus menerus dalam hal kegiatan membaca, sehingga seseorang tersebut paham dengan isi buku yang telah dibaca. Jika seseorang memiliki minat baca yang rendah maka dapat dikatakan cara untuk mendapatkan informasi yang diperoleh dapat melalui media eletronik dan bukan dari buku yang berada di perpustakaan. Minat baca dapat dipengaruhi beberapa faktor yaitu yang berasal dari dalam individu (personal) dan luar individu (institusional). Personal yaitu yang didapat dari dalam diri sendiri yang diantaranya faktor jenis kelamin, usia, kemampuan membaca, intelegensi, kebutuhan psikologis dan sikap. Sedangkan institusional yaitu yang didapat dari luar individu yang diantaranya faktor status sosial, ekonomi dan ketersediaan buku-buku, serta dari teman sebaya, orangtua dan guru. Dengan adanya faktor personal yang ada didalam diri seseorang termasuk usia dapat mempengaruhi bagaimana ketertarikan atau kesukaan mereka terhadap kegiatan membaca. Selain itu dapat mempengaruhi seberapa banyak atau lama mereka menghabiskan waktu dengan berkunjung ke perpustakaan untuk membaca buku, sehingga itu dapat berpengaruh terhadap seberapa banyak orang yang melakukan peminjaman buku di perpustakaan tersebut yang dapat dikatakan jika sering melakukan peminjaman buku di perpustakaan maka memiliki ketertarikan atau kesukaan terhadap kegiatan membaca.

Dari buku yang berjudul Pemetaan Minat Baca Masyarakat, untuk melihat tingkat minat baca dikatakan bahwa setiap reponden dapat menghabiskan waktu dalam membaca rata-rata 1 sampai 2 jam untuk setiap hari nya. Sedangkan masyarakat yang memiliki minat membaca tinggi dapat menghabiskan waktu rata-rata 2 jam lebih untuk setiap harinya. Selain itu disimpulkan pula bahwa untuk usia<12 tahun yaitu kategori anak-anak, usia >12-25 tahun yaitu kategori remaja, usia >25-45 tahun yaitu kategori dewasa, usia >45-65 tahun yaitu kategori lansia, dan usia >65 tahun yaitu kategori manula. Dan tingkat minat baca yang tinggi yaitu pada usia >65 tahun dengan lama membaca 1 sapai 2 jam setiap hari dengan presentase 46,55%.

2. Permasalahan

Berdasarkan Data Prediksi Minat Baca tahun 2017/2019, jumlah prediksi minat baca berdasarkan usia yang diperoleh dari 20 data pengunjung ini sebanyak 14 pengunjung yang minat baca dengan rata-rata termasuk ke dalam kategori umur masa remaja akhir (17 - 25 tahun), dan 6 pengunjung yang tidak minat baca dengan rata-rata termasuk ke dalam kategori umur masa remaja akhir (17-25 tahun).

Pengaruh minat baca berdasarkan usia dari tabel diatas dapat dilihat dari tanggal pinjam, tanggal jatuh tempo dan tanggal kembali. Jika pengunjung melakukan peminjaman buku dan melakukan pengembalian sebelum tanggal jatuh tempo atau tidak lebih dari 10 hari maka dapat dikatakan “Ya” atau minat baca. Sedangkan jika melakukan pengembalian lewat dari tanggal jatuh tempo dan lebih dari 10 hari dapat dikatakan “Tidak” atau tidak minat baca. Sehingga dalam menentukan prediksi minat baca yang dilihat dari data peminjaman masih dilakukan secara manual atau dilakukan secara tertulis. Hal itu akan menghambat bagian petugas perpustakaan untuk melakukan prediksi minat baca. Maka dapat diidentifikasikan masalah, yaitu :

Volume 11 Number 1 May 2021 Page. 67-74
Journal Homepage : http://teknois.stikombinaniaga.ac.id/index.php/JBS
DOI Link : http://doi.org/10.36350/jbs.v11i1

a. Belum akuratnya dalam memprediksi minat baca berdasarkan usia
b. Belum efektifnya proses dalam memprediksi minat baca.

3. Tujuan

Adapun tujuan dari penelitian ini adalah :

a. Menerapkan metode Naive Bayes untuk memprediksi minat baca berdasarkan usia.
b. Mengukur tingkat ketepatan dan tingkat efektivitas metode Naive Bayes dalam memprediksi minat baca berdasarkan usia.

4. Tinjauan Pustaka

a. Data Mining

Menurut larose (2014), dikutip dari buku dengan judul "*Discovering Knowledge Ind Data: An Introduction To data Mining*" menjelaskan bahwa data mining adalah kegiatan menganalisis sejumlah data yang besar. Setelah data diproses, korelas, dan pola baru ditemukan dengan cara yang berbeda dari masa lalu, yaitu penggunaan teknik pengenalan pola dan teknik statistik dan analisis untuk memfilter sejumlah data yang besar yang disimpan dalam data repository. Dengan teknik matematika yang bisa dimengerti dan berguna untuk pemilik data.

Menurut Tan (2016), serangkaian proses dibagi menjadi beberapa tahapan sebagai berikut:

1) Pembersihan data yaitu untuk mengilangkan data dan menghapus data yang tidak sesuai
2) Integrasi data yaitu kombinasi data dari berbagai sumber
3) Transformasi data yaitu mengubah menjadi bentuk yang dapat diproses mining
4) Aplikasi teknik data mining yaitu proses mengestrak pola dari data yang sudah ada
5) Evaluasi pola yang didapat yaitu proses menafsirkan pola sebagai pengetahuan yang digunakan untuk pengambilan pendukung keputusan.
6) Presentasi pengetahuan yaitu dengan memberikan pengetahuan.

b. Clustering

Menurut Larose, Daniel T, larose, Chantal D (2014), dikutip dari buku yang berjudul "*Discovering Knowledge Ind Data: An Introduction To Data Mining*", menjelaskan bahwa clustering adalah kumpulan record yang serupa dan berbeda dari record di cluster lain. *Clustering* mengacu pada pengelompokkan catatan, pengamatan, atau kasus ke dalam kelas benda serupa. Algoritma *clustering* untuk mencari atau mengelompokkan seluruh data yang diatur ke dalam sub kelompok atau *clustering* yang relatif homogeny, dimana kesamaan catatan dalam *clustering* di maksimalkan dan kesamaan untuk catatan diluar cluster ini dikurangi. Clustering biasanya merupakan tahap pertama di dalam proses data mining, sehingga dari cluster tersebut dihasilkan dapat digunakan lebih lanjut untuk berbagai teknik.

B. METODE

1. Algoritma Naïve Bayes

Algoritma naïve bayes adalah metode klasifikasi menggunakan probabilitas dan statistic yang diajukan oleh ilmuwan Inggris Thomas Bayes. Algoritma naïve bayes memprediksi peluang masa depan berdasarkan pengalaman masa lalu, sehingga disebut teorema bayes. Fitur utama dari pengklasifikasian naïve bayes adalah asumsi yang sangat kuat (naif), yang tidak bergantung pada setiap kondisi ata peristiwa.

Menurut Karthika dan sairam (2015), menjelaskan bahwa dengan asumsi klasifikasi ini, metode naïve bayes merupakan metode klasifikasi yang sangat sederhana. Dengan menggguakan metode naïve bayes, cari terlebih dahulu nilai probabilitas dan kemungkinan maksimum dari masing-masing atribut dari setiap kategori.

Persamaan dari probabilitas prior:

$$P(H) = \frac{NJ}{N}$$

Volume 11 Number 1 May 2021 Page. 67-74
Journal Homepage : http://teknois.stikombinaniaga.ac.id/index.php/JBS
DOI Link : http://doi.org/10.36350/jbs.v11i1

Dimana :

Nj : jumlah data pada suatu class

N : jumlah total data

Persamaan dari Teorema Bayes (Saputra et al.,2018) :

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

Dimana :

X : data class yang belum diketahui

H : data hipotesis yaitu suatu class spesifik

P(H|X) : jumlah probabilitas hipotesis H berdasarkan kondisi X (posterior probabilitas)

P(H) : jumlah probabilitas hipotesis H (prior probabilitas)

P(X) : jumlah probabilitas X

P(X|H) : jumlah probabilitas X berdasarkan kondisi pada hipotesis H

Penentuan class dilakukan dengan membandingkan nilai probabilitas dari suatu sampel yang berada di class dengan nilai probabilitas suatu sampel yang berada di class lain. Untuk menentukan class yang tepat dari suatu data sampel dapat melakukan perbandingan nilai posterior untuk setiap masing-masing class, dan membandingkan class dengan nilai class posterior yang tertingggi.

Adapun algoritma penyelesaian dari Metode *Naive Bayes* dapat di lihat pada gambar berikut

Gambar 1. Algoritme Naïve Bayes

2. Teknik Analisa Data

Untuk uji hasil keakuratan dalam penelitian ini menggunakan confusion matrix (*akurasi*). Confusion matrix adalah metode yang digunakan untuk menghitung keakuratan konsep data mining. Akurasi atau keyakinan adalah laporan kasus prediksi positif, dan juga benar-benar positif untuk data actual. Recall atau sensitivitas adalah kategori kasus positif yang benar-benar diprediksi dengan benar menjadi positif.

Tabel 1. Model Confusion Matrix

Aktual	Classified as	
	+	-
+	True Positives	False Negatives
-	False Positives	True Negatives

Volume 11 Number 1 May 2021 Page. 67-74
Journal Homepage : http://teknois.stikombinaniaga.ac.id/index.php/JBS
DOI Link : http://doi.org/10.36350/jbs.v11i1

Rumus yang digunakan untuk confusion matrix adalah :

$$\text{Akurasi} = \frac{TP + TN}{TP+TN+FP+FN} \times 100\%$$

Pada pengukuran kinerja menggunakan confusion matrix, terdapat 4 istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah

a. TP (True Positives) : yaitu jumlah dari data positif namun dihasilkan sebagai data positif.
b. TN (True Negatives) : yaitu jumlah dari data negatif namun dihasilkan sebagai data positif.
c. FP (False Positives) : yaitu jumlah dari data negatif namun dihasilkan sebagai data positif.
d. FN (False Negatives) : yaitu jumlah dari data positif namun terdeteksi sebagai data negatif

C. HASIL DAN PEMBAHASAN

1. Hasil

a. Menentukan Variabel

Dalam memprediksi minat baca berdasarkan usia di dalam penelitian ini menggunakan data peminjaman dan diperoleh 5 variabel yang telah dilakukan preprocessing dari 7 variabel, dengan melakukan pengelompokkan dari data yang didapat. Variabel yang digunakan untuk melakukan prediksi yaitu nama anggota, kategori buku, jumlah hari telat, kelompok umur, dan tingkat minat baca. Pada variabel tingkat minat baca terdapat dua kemungkinan yaitu tinggi dan rendah, dan terdapat 50 data yang didapat dari data peminjaman Perpustakaan GOR Pajajaran. Berikut tabel 4.18 data prediksi yang akan digunakan.

Tabel 2. Data Prediksi

No	Nama	Kriteria 1	Kriteria 2	Kriteria 3	Tingkat Minat Baca
1	Asmiati	Ilmu Sosial	Tidak Tepat Waktu	Masa Remaja Akhir (17-25 tahun)	Rendah
2	Ririn Rinaya	Sejarah dan Geografi	Tidak Tepat Waktu	Masa Remaja Akhir (17-25 tahun)	Tinggi
3	Irma Rita	Ilmu Murni	Tidak Tepat Waktu	Masa Dewasa Akhir (36-45 Tahun)	Tinggi
4	Raya Surajat	Ilmu Murni	Tidak Tepat Waktu	Masa Remaja Awal (12-16 tahun)	Rendah
5	Syarif Hidayat	Agama	Tepat Waktu	Masa Remaja Akhir (17-25 tahun)	Tinggi
…	…	…	…	…	…
50	Dwi Faza	Ilmu Sosial	Tidak Tepat Waktu	Masa Dewasa Akhir (36-45 Tahun)	Tinggi

Volume 11 Number 1 May 2021 Page. 67-74
Journal Homepage : http://teknois.stikombinaniaga.ac.id/index.php/JBS
DOI Link : http://doi.org/10.36350/jbs.v11i1

Berikut keterangan dari setiap variabel dapat dilihat pada tabel 3.

Tabel 3. Keterangan Variabel

Variabel	Kriteria
Kategori Buku	Kriteria 1
Jumlah Hari Telat	Kriteria 2
Kelompok Umur	Kriteria 3

Berikut keterangan dari variabel jumlah hari telat pada tabel 4.

Tabel 4. Keterangan Jumlah Hari Telat

Jumlah Hari telat	Keterangan
Tepat Waktu	< 7 Hari
Tidak Tepat Waktu	> 7 Hari

Dari data tersebut diperoleh data uji yang nantinya akan di terapkan ke dalam metode naive bayes. Berikut data uji dapat dilihat pada tabel 5.

Tabel 5. Data Uji

No	**Nama**	**Kriteria 1**	**Kriteria 2**	**Kriteria 3**	**Tingkat Minat Baca**
1	Ulfah	Ilmu Murni	Tepat Waktu	Masa Remaja Akhir (17-25 tahun)	?

b. Menghitung Nilai Peluang (probabilitas) dari kasus baru setiap kelas (label) yang ada "P(XK|Ci)"

Tahap pertama yang dilakukan yaitu menghitung probabilitas total masing-masing kelas kejadian atau dari 7 variabel yang digunakan. Terlebih dahulu menentukan peluang dari variabel tingkat minat baca yaitu "tinggi" dan "rendah". Caranya adalah dengan menentukan peluang dari kasus yang terjadi dengan setiap variabel di data peminjaman, membagi jumlah data kelas kejadian dengan jumlah seluruh data di tabel.

Maka perhitungannya dapat menjadi sebagai berikut

1) P (Y=Tinggi) = 25/50 atau jumlah data peminjaman buku "Tinggi" pada kejadian "tingkat minat baca" dibagi jumlah seluruh data
2) P (Y= Rendah) = 25/50 atau jumlah data peminjaman buku "Rendah" pada kejadian "tingkat minat baca" dibagi jumlah seluruh data
3) P (Kriteria 1 = "Ilmu Murni" | TIngkat = "Tinggi") P = 3/ 25 = 0,08
4) P (Kriteria 2 = "Tepat Waktu" | Tingkat = "Tinggi") P = 17 / 25 = 0,7
5) P (Kriteria 3 = "Masa Remaja Akhir (17-25 tahun)" | Tingkat = "Tinggi") P = 7 / 25 = 0,3
6) P (Kriteria 1 = "Ilmu Murni" | Tingkat = "Rendah") P = 3 / 25 = 0,12
7) P (Kriteria 2 = "Tepat Waktu" | Tingkat = "Rendah") P = 6 / 25 = 0,24
8) P (Kriteria 2 = "Masa Remaja Akhir 17-25 tahun" | Tingkat = "Rendah") P = 6 / 25 = 0,24

Volume 11 Number 1 May 2021 Page. 67-74
Journal Homepage : http://teknois.stikombinaniaga.ac.id/index.php/JBS
DOI Link : http://doi.org/10.36350/jbs.v11i1

c. Menghitung nilai akumulasi peluang dari setiap kelas (label) "P(X|Ci)"

Tahap kedua adalah menghitung probabilitas setiap kasus. Perhitungan dilakukan dengan menghitung jumlah kasus yang terjadi di masing-masing variabel, sesuai yang bersangkutan dengan data tambahan, dengan masing- masing kelas kejadian.

Maka perhitungannya adalah sebagai berikut:

1) P (Tingkat Minat Baca = "Tinggi") x P (Y = "Tinggi") = 0,08 x 0,7 x 0,3 = 0,015232
2) P (Tingkat Minat Baca = "Rendah") x P (Y ="Rendah") = 0,12 x 0,24 x 0,24 = 0,006912

d. Menghitung nilai probabilitas akhir setiap kelas (label) "P(X|Ci) x P(Ci)"

Tahap ketiga adalah mengalikan semua hasil variabel pada setiap kelas kejadian. Maka perhitungannya adalah sebagai berikut :

1) Kelas kejadian tingkat minat baca "Tinggi"

   P (X | Tingkat Minat Baca = "Tinggi") x (X | Tingkat Minat Baca = "Tinggi")

   0,015232 x 25 / 50 = 0,007616

2) Kelas kejadian tingkat minat baca "Rendah"

   P (X | Tingkat Minat Baca = "Rendah") x (X | Tingkat Minat Baca = "Rendah")

   0,006912 x 25 / 50 = 0,003456

e. Menentukan nilai probabilitas akhir terbesar dari setiap kelas

Pada tahap terakhir ini, yang dilakukan hanya membandingkan hasil akhir dari setiap kelas. Hasil atau keputusan yang diambil adalah hasil yang paling besar. Untuk data diatas, maka hasilnya adalah :

Tabel 6. Hasil data uji

No	Nama	Kriteria 1	Kriteria 2	Kriteria 3	Tingkat Minat Baca
1	Ulfah	Ilmu Murni	Tepat Waktu	Masa Remaja Akhir (17-25 tahun)	Tinggi

Sehingga dengan *Naive Bayes* dapat disimpulkan **"Tinggi"** untuk data input ini, berdasarkan estimasi probabilitas yang dipelajari dari data training. Dengan normalisasi, agar jumlah probabilitas sama dengan 1, bisa menghitung *probabilitas conditional* untuk pilihan "Tingkat Minat Baca Sedang" jika diberikan nilai-nilai atribut. Untuk contoh ini, probabilitasnya adalah :

$$\textbf{Probabilitas} = \frac{0{,}007616}{0{,}007616+0{,}003456} = 1{,}003456$$

2. Pembahasan

Pada tahap ini dilakukan pengukuran keakuratan antara hasil yang diperoleh menggunakan confusion matrix. Pengukuran aturan dilakukan dengan membandingkan hasil prediksi data latih berdasarkan variable yang telah ditentukan dengan data yang seharusnya yang merupakan data nyata. Hasil pengukuran tersebut dapat dilihat pada tabel 7.

Volume 11 Number 1 May 2021 Page. 67-74
Journal Homepage : http://teknois.stikombinaniaga.ac.id/index.php/JBS
DOI Link : http://doi.org/10.36350/jbs.v11i1

Tabel 7. Hasil Confusion Matrix

Aktual Class	Predicted as	
	Tinggi	Rendah
Tinggi	5	1
Rendah	1	3

Berdasarkan tabel 7 maka dapat dilakukan perhitungan akurasi sebagai berikut :

$$\text{Akurasi} = \frac{TP + TN}{TP+TN+FP+FN} \times 100\%$$

$$\text{Akurasi} = \frac{5+3}{5+3+1+1} \times 100\%$$

$$\text{Akurasi} = \frac{8}{10} \text{ x } 100\% = 80\%$$

D. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, kesimpulan yang bisa diuraikan bahwa Dengan menerapkan metode Naive Bayes dapat dijadikan sebagai sistem pendukung keputusan untuk melakukan prediksi minat baca berdasarkan usia. Dilakukan pengukuran nilai akurasi dalam penerapan metode Naive Bayes untuk uji hasil dengan hasil 80% yang disimpulkan layak dan bagus untuk digunakan.

E. DAFTAR PUSTAKA

[1] Armstrong, J.S.. "*Relative Accuracy of Judgemental and Extrapolative Methods in Forecasting Annual Earnings*". *Journal ofForecasting* No.02 (1983): 437-447.

[2] Budi Santosa. (2007). *Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bisnis*. Yogyakarta

[3] Departemen Pendidikan Nasional, Perpustakaan Nasional Republik Indonesia. (2007). Pemetaan Minat Baca Masyarakat. Di Tiga Provinsi: Sulawesi Selatan, Riau dan Kalimantan Selatan

[4] Eko Prasetyo. (2014). Data Mining Konsep dan Aplikasi Menggunakan Matlab. Yogyakarta

[5] Ghaniy, Rajib, and Kepin Sihotang. "Penerapan Metode Naïve Bayes Classifier untuk Penentuan Topik Tugas Akhir." Teknois, vol. 9, no. 1, 16 May. 2019, pp. 63-72, doi:10.36350/jbs.v9i1.7.

[6] Jogiyanto, Hartono, 2005. *Analisis & Desain Sistem Informasi Pendekatan Terstruktur Teori dan Praktek Aplikasi Bisnis*, Yogyakarta: Andi Offset

[7] Karthika, S., & Sairam, N. (2015). A Naïve Bayesian Classifier for Educational Qualification. Indian Journal of Science and Technology, 8(16), 1–5. http://doi.org/10.17485/ijst/2015/v8i16/62055

[8] Larasati Milburga, et.all. 1986. "Membina Perpustakaan Sekolah"Yogyakarta:Kanisius, hlm. 17

[9] Larose, Daniel T, Larose, Chantal D. (2014). Discovering Knowledge Ind Data: An Introduction To Data Mining – Second Edition, John Willey & Sons. Inc., Hoboken, New Jersey

[10] Tan, P.N., Steinbach, M.,Kumar,V.(2006).Introduction to Data mining. Boston:Pearson Education