

Sistem Pendukung Pengambilan Keputusan Kesesuaian Beasiswa Menggunakan Naive Bayes

Tan, Yohanes Christianto Aryo Wicaksono

Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Soegijapranata
christ.yohanes@gmail.com

YB Dwi Setianto

Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Soegijapranata
setianto@unika.ac.id

Abstract

Until now there are no system that helps determine candidates who are eligible to receive the Sandjojo Foundation scholarship at Soegijapranata Catholic University. Therefore, scholarship recipients are sometimes not on target. This research can help solve problems by using the Naive Bayes algorithm as a medium for decision makers with criteria such as based on the student's GPA, the total income of parents, and must be an active student in the organization. Using previous registrant data, the Naive Bayes algorithm can study the data for each of the criteria entered to determine whether applicants qualify or not to receive scholarships from the Sandjojo Foundation. This research was made using a web-based application and has been tested four (4) times with a variety of test and training data, with an accuracy of around 50% to 65% and a time of less than three (3) seconds.

Keywords: naive bayes, scholarship, decision support system.

Abstrak

Hingga saat ini belum ada sistem yang membantu menentukan kandidat yang layak menerima beasiswa Yayasan Sandjojo di Universitas Katolik Soegijapranata. Oleh karena itu, penerima beasiswa kadang tidak tepat sasaran. Penelitian ini dapat membantu menyelesaikan masalah dengan menggunakan algoritma Naive Bayes sebagai media bagi para pengambil keputusan dengan kriteria seperti berdasarkan IPK pendaftar siswa, total pendapatan orang tua, dan harus seorang siswa aktif dalam organisasi. Dengan menggunakan data pendaftar sebelumnya, algoritma Naive Bayes dapat mempelajari data untuk setiap kriteria yang dimasukkan untuk menentukan apakah pelamar memenuhi syarat atau tidak untuk menerima beasiswa dari Sandjojo Foundation. Penelitian ini dibuat menggunakan aplikasi berbasis web dan telah diuji empat (4) kali dengan berbagai macam data uji dan data pelatihan, dengan hasil akurasi sekitar 50% hingga 65% dan waktu lebih rendah dari tiga (3) detik.

Kata kunci: naive bayes, scholarship, decision support system.

PENDAHULUAN

Penentuan Penerima beasiswa oleh UPT bidang 3 Unika Soegijapranata sampai saat ini masih dilakukan secara manual, hal ini mengakibatkan proses pengambilan keputusan tidak dapat dilakukan dengan cepat, akurat, dan efisien. Sering terjadi penentuan penerima beasiswa yang kurang tepat sasaran, untuk itu projek ini diyakini akan membantu otomatisasi pengambilan keputusan kelayakan penerima beasiswa Yayasan Sandjojo.

Dengan keluaran projek ini diyakini dapat menambah ketepatan penerima beasiswa Yayasan Sandjojo. Menggunakan metode Naive Bayes penentuan penerima beasiswa Yayasan Sandjojo dapat dipermudah karena metode mempelajari data penerima beasiswa yang sudah ada. Kemudian dikonversi dalam susunan suatu array 2 dimensi yang diolah dan menghasilkan nilai probabilitas sebagai penentuan class tiap data.

Penelitian ini akan mengelompokan data menjadi 2 kelas yaitu mahasiswa layak dan tidak layak untuk menerima beasiswa “Yayasan Sandjojo” dengan faktor penentu IPK, penghasilan orang tua, dan keaktifan mahasiswa dalam organisasi.

TELAAH LITERATUR

Dalam menentukan kelayakan mahasiswa sebagai mahasiswa yang dianggap layak menerima beasiswa adalah suatu hal yang bisa dibilang sepele hal itu merupakan hal yang penting.

Dalam jurnal yang berjudul “Algoritma Naive Bayes Untuk Mencari Perkiraan Waktu Studi Mahasiswa” (Jananto, 2013) menggunakan metode Naive Bayes untuk menentukan lama studi mahasiswa berdasarkan IPK, jenis kelamin, kota lahir, tipe sekolah, kota sekolah. Arief Jananto melakukan ini dengan step, Mining data, Konversi data untuk tiap faktor, dan yang terakhir implementasi Naive Bayes. Hasil yang diperoleh dari 66 data tingkat kesalahan 20 record dengan presentase 66% benar. seperti yang dilakukan Arief Jananto projek ini memiliki kesamaan dalam hal parameter IPK sebagai salah satu faktor penentu dalam perhitungan Naive Bayes.

Jurnal “METODE NAIVE BAYES UNTUK PENENTUAN PENERIMA BEASISWA BIDIKMISI UNIVERSITAS MULAWARMAN” (Dahri, Agus & Khairi, 2016) Penerima beasiswa bidikmisi di Universitas Mulawarman sering kali kurang tepat sasaran dan memakan waktu lambat. Penerima beasiswa menjadi tidak konsisten penyeleksipun tidak transparan pada penerima, dibuatlah Sistem Pendukung

Keputusan (SPK) untuk menyeleksi. SPK didalam jurnal ini ada beberapa komponen. Data Mangement sebagai Database, Model Management untuk perhitungan, Communication sebagai media interface, dan Knowledge Management. Hasil yang didapatkan akurasi prediksi benar dari SPK mencapai 85,56% dari 90 data uji dari jumlah data training 278 data. SPK jurnal ini menggunakan parameter penghasilan orang tua yang sama dengan penelitian ini. Penelitian ini menggunakan data uji sebanyak 368 data training.

Berdasarkan dari jurnal Mantik Penusa dengan judul “PENERAPAN ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI JUMLAH PRODUKSI BARANG BERDASARKAN DATA PERSEDIAAN DAN JUMLAH PEMESANAN PADA CV. PAPADAN MAMA PASTRIES” (Manalu, Sianturi, & Manalu, 2017) dalam produksi roti dapat kelebihan dan kekurangan stok roti. Effrida Manalu, et al membuat sistem yang mudah digunakan, akurat, efisien dan hemat waktu dengan menerapkan metode Naive Bayes dengan tujuan dapat mengantisipasi kelebihan atau kekurangan produksi. Effrida Manalu, et al (2017:17) menyimpulkan “Metode ini membutuhkan data traning yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian.”. Tahapan baca data

training, hitung, menentukan value terbesar untuk dijadikan kesimpulan. Algoritma Naive Bayes membantu memprediksi produksi stok roti. Penelitian ini memiliki kesamaan software penyimpanan data MySQL. Perbedaannya penelitian ini hanya ada 1 macam data mahasiswa dalam jurnal terdapat 3 macam data roti tawar kupas, roti tawar gandum, dan roti tawar cokelat.

Pada jurnal “SISTEM PENDUKUNG KEPUTUSAN KLASIFIKASI TINGKAT KEGANASAN KANKER PAYUDARA DENGAN METODE NAIVE BAYES CLASSIFIER” (Via, Nugroho, & Syafrizal, 2015) Kanker Payudara salah satu jenis kanker yang sering ditemukan pada kebanyakan wanita. Yisti Vita Via, et al membuat keluaran sistem untuk mengklasifikasi tingkat keganasan kanker payudara dengan metode Naive Bayes Classifier. Karena terbukti memiliki akurasi dan kecepatan yang tinggi. Tahapan yang digunakan definisi masalah, pengumpulan data, pengolahan data, menentukan alternatif solusi. Yisti Vita Via, et al bertujuan untuk mengklasifikasi keganasan kanker payudara dengan hasil akurasi 97,82% benar dan 2,18% salah dengan 137 data test. Persamaan penelitian ini mengklasifikasikan 2 class. Perbedaan dengan penelitian ini menggunakan 3 attribut, Yisti Vita Via, et al menggunakan 9 attribut.

Berdasarkan Jurnal “Sistem Pendukung Keputusan Penentuan Kelayakan Calon Tenaga Kerja Indonesia Menggunakan Metode Naive Bayes” (Wasiati & Wijayanti, 2014) Perusahaan PT Karya Tama membutuhkan tenaga kerja yang layak. Pada tahun 2012/2013 terdapat 542 pendaftar. 385 diterima. Hera Wasiati dan Dwi Wijayanti merancang suatu aplikasi berbasis dekstop penentu kelayakan tenaga kerja dengan metode Naive Bayes. Kriteria penentu usia, pendidikan, tinggi badan, berat badan, dan hasil test. Hasil perancangan aplikasi penentu kelayakan tenaga kerja perusahaan dapat menyeleksi pendaftar dengan mudah dan cepat dengan akurasi benar 73,89% benar dengan 180 data uji dan 362 data training. Persamaan penelitian ini pada tampilan aplikasi menu login, menu akurasi, menu input pendaftar. Perbedaan penelitian ini berbasis website.

Berdasarkan Aritikel Ilmiah “Implementasi Algoritma Naive Bayes dalam Proses Analisis Efisiensi Jasa Servis Bengkel Daihatsu Cabang Salatiga” (Jatmikanto & Pakereng, 2016) Untuk mencapai efisiensi jasa servis di bengkel Daihatsu cabang Salatiga artikel ini menggunakan algoritma Naive Bayes dan teknik Data Mining untuk mendapatkan informasi keluhan di bengkel tersebut. Jatmikanto mengolah riwayat data dengan data mining, wawancara ke customer,

merancang aplikasi, uji coba. Hasil dari artikel ini membagi menjadi 3 class dengan akurasi keberhasilan 80% dan kesalahan 20% dari 10 data uji dengan 684 data training. Artikel ini sama dengan penelitian ini dengan terdapatnya keluaran berupa aplikasi berbasis Desktop, artikel ini menggunakan database SQL Server Express 2008 sebagai database.

Berdasarkan artikel dalam jurnal “Sistem Persediaan Obat pada Puskesmas menggunakan Metode Naive Bayes (Studi Kasus Puskesmas Talang Tinggi Seluma)” (Furqan, Wibowo & Abdullah, 2016) Puskesmas Talang Tinggi masih menggunakan sistem yang bersifat manual, ketika memesan obat harus melakukan rekap satu persatu, hal ini akan menyebabkan kekeliruan karena terdapat 200 jenis obat. Furqan et al. Merancang sistem prediksi persediaan sebelum obat itu habis yang diolah menggunakan metode Naive bayes dengan harapan mudah digunakan. Furqan et al. Merancang sistem prediksi persediaan stok obat dengan survey, analisa, desain, persiapan hardware dan software, implementasi sistem, Pemeliharaan sistem. Furqan et al. Menggunakan 3 parameter sebagai penentu yaitu satuan obat, golongan obat, dan jumlah obat. Hasil yang didapatkan dalam jurnal ini membantu proses Prediksi obat. Penelitian ini sama menggunakan 3 attribute sebagai penentu. Perbedaan

penelitian ini pada bahasa pemrograman yang digunakan oleh jurnal yaitu Visual Basic 6.0. Berdasarkan “Sistem Prediksi Kesiapan Siswam dalam Menghadapi Ujian Nasional menggunakan Metode Naive Bayes” (Putri & Irsyadi, 2018) Sekolah tidak mengetahui siswa yang kurang siap untuk menghadapi ujian nasional yang menjadi kewajiban tiap siswa untuk dinyatakan lulus sekolah. Putri & Irsyadi merancang Sistem Prediksi Kesiapan Siswa SMP dalam Menghadapi Ujian Nasional Menggunakan Metode Naive Bayes dengan tahapan pengumpulan data dengan data mining, pengolahan data perhitungan Theorema Bayes. Hasil jurnal ini membantu memprediksi kesiapan siswa menghadapi ujian nasional dengan akurasi 93,35% benar dengan percobaan sebanyak 6 kali dengan data uji yang berbeda, menggunakan kurang lebih 145 data training setiap test dan kurang lebih 130 data test. Penelitian ini sama menggunakan tampilan website. Perbedaan dengan penelitian ini tidak menggunakan metode data mining untuk pengumpulan data training.

Jadi projek ini menghasilkan keluaran aplikasi berbasis website dengan menggunakan metode Naive Bayes seperti yang dilakukan Jananto, 2013; Dahri et al., 2016; Manalu et al., 2017; Via et al., 2015; Wasiati & Wijayanti, 2014; Jatmikanto &

Pakereng 2016; Furqan et al., 2016; Putri & Irsyadi, 2018; untuk menentukan kelayakan penerima beasiswa Yayasan Sandjojo dengan ketentuan IPK, penghasilan orang tua, dan Keatifan mahasiswa dalam organisasi seperti halnya atribut yang digunakan sebagai penentu. Dengan 368 data training dan 114 data uji yang didapat dari UPT bidang 3.

METODE PENELITIAN

Dalam pembuatan penelitian ini langkah – langkah yang harus dilalui untuk mendapatkan hasil yang diinginkan:

a. Pengumpulan Data

Data didapatkan secara langsung bukan dari hasil wawancara atau menyebarkan pertanyaan. Meminta secara langsung kepada UPT bidang 3 di Universitas Katolik Soegijapranata yang mengurus beasiswa Yayasan Sandjojo. Penelitian ini menggunakan 2 periode pendaftaran yaitu 2016 dan 2017.

b. Pengumpulan dan Analisa Jurnal

Dalam pembuatan penelitian ini membutuhkan landasan teori untuk memperkuat pembuatan penelitian ini. Jurnal – jurnal diambil dari website <https://scholar.google.co.id/> dan terlampir pada bagian daftar pustaka. Penelitian ini menggunakan 8 jurnal sebagai referensi.

c. Desain dan Implementasi

Tahap yang dilalui untuk mencapai tujuan penelitian ini adalah penerjemahan kedalam bahasa pemrograman. Design tampilan untuk mempermudah pengguna melihat data. Penelitian ini menggunakan bahasa pemrograman PHP untuk membuat sistem pendukung pengambil keputusan dengan metode Naive Bayes yang memerlukan database. Penelitian ini memerlukan Myql untuk menyimpan data training.

d. Uji Coba

Testing untuk menguji akurasi dari program dari data test. Penelitian ini menggunakan 3 metode test antara lain test fungsi, test performa, dan test akurasi.

e. Membuat Laporan

Pembuatan laporan ini sebagai berkas untuk menjelaskan langkah – langkah dalam mengerjakan penelitian, sekaligus untuk laporan dalam menyelesaikan penelitian akhir., maka tidak perlu disebutkan sumbernya.

HASIL DAN PEMBAHASAN

Penelitian ini dibuat menggunakan bahasa pemrograman PHP karena PHP aplikasi di universitas Katolik Soegijaprana kebanyakan menggunakan aplikasi berbasis website dengan pemrograman PHP. Penelitian ini menggunakan database Mysql untuk

menyimpan data training dan data uji, data dalam Mysql mudah untuk diolah karena akan dimasukan dalam struktur data Array 2 dimensi dan server di Universitas Katolik Soegijaprana menggunakan Mysql sebagai penyimpanan data di aplikasi Sintak. Penelitian ini menggunakan algoritma Naive Bayes karena Thomas Bayes sebagai ilmuwan penemu Theorema Bayes mengemukakan memprediksi peluang dimasa depan berdasarkan pengalaman di masa sebelumnya.

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

Keterangan :

X = Data dengan class yang belum diketahui

H = Hipotesis data X merupakan suatu class spesifik

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi X (posteriori prob.)

P(H) = Probabilitas hipotesis H (prior prob.)

P(X|H) = Probabilitas X berdasarkan kondisi tersebut

P(X) = Probabilitas dari X

Naive Bayes dalam penelitian ini menggunakan 3 parameter yaitu IPK, penghasilan orang tua, dan keaktifan mahasiswa dalam organisasi. Metode Naive Bayes akan membagi data menjadi 2 cluster yaitu layak dan tidak layak dengan penambahan peringkat berdasarkan

perhitungan hasil value layak. keputusan apabila hasil dari data input value layak semakin tinggi maka data tersebut akan muncul paling awal sebagai data prioritas. Penelitian ini menggunakan range IPK sebagai berikut :

Tabel 1. Pengelompokan Berdasarkan IPK

Kategori	Minimal	Maximal
I1	0	2.75
I2	2.76	3.25
I3	3.26	2.50
I4	3.51	4.00

Pembagian IPK dimaksudkan untuk mengelompokan data training dan data uji. Penghasilan orang tua di kelompokan berdasarkan total gaji, pengelompokan total penghasilan orang tua yaitu :

Tabel 2. Pengelompokan Berdasarkan Total Penghasilan Orang Tua

Kategori	Minimal	Maximal
P1	0	999.999
P2	1.000.000	2.899.999
P3	2.900.000	7.899.999
P4	7.900.000	~

Gaji orang tua dihitung dengan penjumlahan gaji ayah dan gaji ibu hasil inputan dari pendaftar beasiswa dengan rumus sebagai berikut :

$$\text{Gaji Total} = \text{Gaji Ayah} + \text{Gaji Ibu}$$

Tabel 3. Pengelompokan berdasarkan jumlah sertifikat

Kategori	Keterangan	Jumlah
S1	Tidak ada	0
S2	Ada sedikit	1 – 6
S3	Ada banyak	> 7

Penelitian ini akan menghitung setiap file sertifikat yang masuk untuk diperhitungkan jumlah sertifikat setiap data. Perhitungan setiap attribute akan memepengaruhi hasil Theorema Bayes. Hasil dari setiap attribute dari setiap kelas akan dikalikan untuk mendapatkan hasil akhir yang nantinya menjadi penentu keputusan data masuk ke class apa layak atau tidak layak. Contoh perhitungan dengan kasus diberikut :

Tabel 4. Contoh data

IPK	Gaji Ayah	Gaji Ibu	Jumlah sertifikat
3.84	Rp. 2.000.000	Rp. 1.500.000	8

Jumlah database : 368 Data

Jumlah database dengan kelas Diterima : 144

Jumlah database dengan kelas Ditolak : 224

$$\text{Total penghasilan} = \text{Gaji Ayah} + \text{Gaji Ibu} = 2.000.000 + 1.500.000 = 3.500.000$$

Perhitungan setiap parameter di kelas Layak (Diterima)

$$P(I4 | \text{Layak}) = 39 / 144 = 0,2708333333$$

$$P(P3 | \text{Layak}) = 53 / 144 = 0,3680555556$$

$$P(S3 | \text{Layak}) = 11 / 144 = 0,0763888889$$

Perhitungan setiap parameter di kelas Tidak Layak (Ditolak)

$$P(I4 | \text{Tidak Layak}) = 50 / 224 = 0,2232142857$$

$$P(P3 | \text{Tidak Layak}) = 122 / 224 = 0,5$$

$$P(S3 | \text{Tidak Layak}) = 12 / 244 = 0,0491803279$$

Perhitungan peluang setiap kelas

$$P(\text{Layak}) = P(I4 | \text{Layak}) \times P(P3 | \text{Layak}) \times P(S3 | \text{Layak})$$

$$P(\text{Layak}) = 0,2708333333 \times 0,3680555556 \times 0,0763888889$$

$$P(\text{Layak}) = 0,0076145753$$

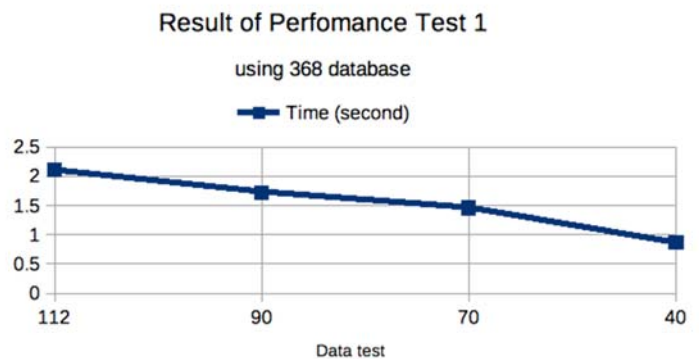
$$P(\text{Tidak Layak}) = P(I4 | \text{Tidak Layak}) \times P(P3 | \text{Tidak Layak}) \times P(S3 | \text{Tidak Layak})$$

$$P(\text{Tidak Layak}) = 0,2232142857 \times 0,5 \times 0,0491803279$$

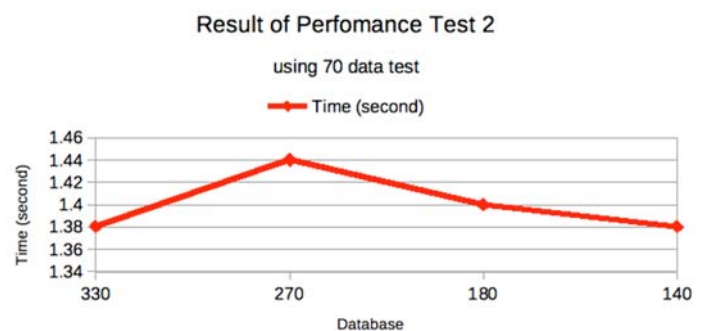
$$P(\text{Tidak Layak}) = 0,0054888759$$

Perhitungan probabilitas tiap kelas menunjukkan hasil probabilitas kelas “Layak” lebih besar dari probabilitas kelas “Tidak Layak” jadi keputusan data ini masuk ke kelas “Layak” Menerima Beasiswa Yayasan Sandjojo.

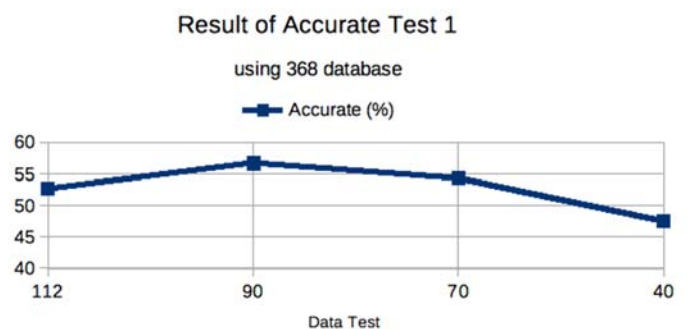
Hasil uji performa dan uji akurasi setiap uji yang dilakukan sebanyak 4 kali dengan kasus data yang berbeda dan hasil yang berbeda pada setiap uji performa dan uji akurasi sebagai berikut :



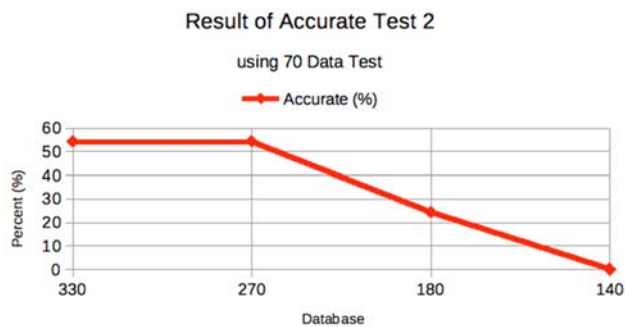
Gambar 1. Uji performa 1



Gambar 2. Uji performa 2



Gambar 3. Uji akurasi 1



Gambar 4. Uji akurasi 2

Dari gambar grafik diatas dapat dilihat kecepatan dalam menentukan kelayak penerima beasiswa dengan metode Naive Bayes kurang dari 3 detik dan akurasi yang dihasilkan mencapai 60% - 65%.

SIMPULAN

Dari hasil implementasi dan pengujian dapat disimpulkan hal – hal sebagai berikut:

1. Dalam projek ini penentu kelayakan penerima beasiswa Yayasan Sandjojo dalam tampilan sudah sesuai dengan rancangan design awal yang di buat.
2. Berdasarkan 4 pengujian performance dapat disimpulkan bahwa Metode Naive Bayes dapat menentukan kelayakan penerima beasiswa kurang dari 2.5 detik. Pengujian akurasi Metode Naive Bayes dalam menentukan kelayakan penerima beasiswa Yayasan Sandjojo lebih dari 50% dengan 4 pengujian.

Saran yang diyakini dapat meningkatkan akurasi dalam pengambil keputusan dengan metode Naive Bayes antara lain :

1. Penambahan kriteria
2. Penambahan database / data training

DAFTAR PUSTAKA

- Dahri, D., Agus, F., & Khairina, D. M. (2017). Metode Naive Bayes Untuk Penentuan Penerima Beasiswa Bidikmisi Universitas Mulawarman. *Jurnal Informatika Mulawarman*, (online), 11(2), 29-36. (<http://e-journals.unmul.ac.id>, diakses 12 September 2018).
- Furqan, M., Wibowo, S. H., Kom, S., & Kom, M. SISTEM PERSEDIAAN OBAT PADA PUSKESMAS MENGGUNAKAN METODE NAÏVE BAYES (STUDI KASUS PUSKESMAS TALANG TINGGI SELUMA), *Jurnal Teknik Informatika Universitas Muhammadiyah Bengkulu*, (online), 1-10. (<https://caridokumen.com/>, diakses 14 September 2018).
- Jananto, A. (2013). Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa. *Dinamik-Jurnal Teknologi Informasi*, (online), 18(1), 9-16. (<https://www.unisbank.ac.id>, diakses 14 September 2018)
- Jatmikanto, T. (2016). Implementasi Algoritma Naïve Bayes dalam Proses Analisis Efisiensi Jasa Servis Bengkel Daihatsu Cabang Salatiga (Doctoral dissertation, *Program Studi Teknik Informatika FTI-UKSW*), (online), (<http://repository.uksw.edu>, diakses 17 September 2018).

- Manalu, E., Sianturi, F. A., & Manalu, M. R. (2017). Penerapan Algoritma Naive Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan dan Jumlah Pemesanan pada Cv. Papadan Mama Pastries. *Jurnal Mantik Penusa*, (online), 1(2), 16-21. (<http://ejournal.pelitanusantara.ac.id>, diakses 10 September 2018).
- Putri, M. P., & Al Irsyadi, F. Y. (2018). Sistem Prediksi Kesiapan Siswa SMP Dalam Menghadapi Ujian Nasional Menggunakan Metode Naive Bayes. *e-Jurnal Mitra Pendidikan*, (online), 2(5), 469-481. (<https://ejournalmitrapendidikan.com>, diakses 10 September 2018).
- Via, Y. V., Nugroho, B., & Syafrizal, A. (2015). Sistem Pendukung Keputusan Klasifikasi Tingkat Keganasan Kanker Payudara dengan Metode Naive Bayes Classifier. *SCAN-Jurnal Teknologi Informasi dan Komunikasi*, (online), 10(2), 63-68. (<http://www.ejournal.upnjatim.ac.id>, diakses 10 September 2018).
- Wasiati, H., & Wijayanti, D. (2014). Sistem Pendukung Keputusan Penentuan Kelayakan Calon Tenaga Kerja Indonesia Menggunakan Metode Naive Bayes (Studi Kasus: Di PT Karyatama Mitra Sejati Yogyakarta). *IJNS-Indonesian Journal on Networking and Security*, (online), 3(2), 45-51. (<http://ijns.org>, diakses 8 September 2018).