

IMPLEMENTASI ALGORITMA *NAÏVE BAYES CLASSIFIER* UNTUK PENILAIAN KINERJA DOSEN

Ahmad Kausar¹, Agus Irawan², Wahyuddin³, Iqbal Fernando⁴

^{1,3}Prodi Teknik Informatika, ²Prodi Sistem Komputer, ⁴Prodi Informatika

^{1,2,3}Fakultas Teknologi Informasi Universitas Serang Raya

⁴Fakultas Sains & Teknik Universitas Falatehan

Jl. Raya Serang Cilegon KM. 5 Kp. Drangong Taktakan Serang

E-mail: akautsar@gmail.com¹, *agusirawan@unsera.ac.id², wahyuddin@unsera.ac.id³, iqbal.28nando@gmail.com⁴

Abstrak - Dosen merupakan tenaga profesional yang bertugas melaksanakan tridharma perguruan tinggi. Dosen Universitas Serang Raya harus selalu mampu memenuhi kewajiban sebagai dosen dengan melaksanakan tridharma perguruan tinggi sebagai bukti kinerja. Pengukuran kinerja dosen sering menimbulkan permasalahan mulai dari objektifitas penilaian yang diragukan, kurangnya tenaga yang memiliki kompetensi untuk mengukur kinerja dosen, sehingga dibutuhkan sebuah solusi untuk menyelesaikan masalah tersebut. Algoritma *Naïve Bayes Classifier* adalah metode untuk mengklasifikasikan data kedalam kelas-kelas berdasarkan probabilitas. *Naïve Bayes Classifier* menggunakan *data training* seperti data kinerja dosen, kemudian untuk mengukur akurasi dibuat *data test*. Pengukuran kinerja dosen di Universitas Serang Raya dalam penelitian ini menggunakan kelas baik sekali, baik, cukup dan kurang, dengan menghasilkan tingkat akurasi sebesar 91% dan error sebesar 8 %.

Kata Kunci: *Classifier*, Kinerja dosen, *Naïve Bayes*

I. PENDAHULUAN

Dosen merupakan elemen penting di perguruan tinggi. Keberhasilan suatu proses pembelajaran dalam perguruan tinggi tidak terlepas dari pengaruh kinerja dosen di perguruan tinggi tersebut. Kinerja dosen merupakan hal yang senantiasa dievaluasi tiap semester di perguruan, tinggi termasuk di Universitas Serang Raya. Hal ini menjadi penting karena bagian tugas dari penyelenggara pendidikan tinggi untuk menjalankan Tri Dharma perguruan tinggi yang terdiri dari aspek : 1) Pendidikan ; 2) Penelitian; dan 3) Pengabdian kepada masyarakat. Setiap semester Dosen diwajibkan untuk melaksanakan Tri Dharma perguruan tinggi, sebagai bagian dari instrument pengukuran kinerja Dosen tersebut.

Di Universitas Serang Raya memiliki lembaga yang mengawasi kebijakan dan mengukur keberhasilan dari suatu kebijakan dijalankan sesuai dengan aturan, lembaga tersebut bernama Badan Penjaminan Mutu Internal atau disingkat BPMI. Salah satu tugas dari BPMI adalah mengawasi jalannya proses akademik, apakah sudah berjalan sesuai dengan aturan universitas atau tidak, termasuk salah satunya mengukur kinerja dosen di setiap semester untuk kemudian dilaporkan kepada pimpinan universitas dalam hal ini Rektor. Data pengukuran kinerja dosen ini diukur dari berbagai instrument yang tersedia, untuk mengukur kinerja pendidikan maka aspek yang akan dilihat adalah presensi dosen mengajar dan kuisioner yang diambil dari seluruh mahasiswa yang mengikuti

perkuliahan dalam semester tersebut. Pengukuran kinerja penelitian dan pengabdian kepada masyarakat didasarkan pada rekapitulasi data penelitian dan pengabdian masyarakat yang dilaporkan oleh setiap dosen dalam kurun waktu satu semester.

Pengukuran ini membutuhkan waktu untuk kemudian di hasilkan laporan kinerja dosen, karena data tersebut di olah tim BPMI sendiri sehingga membutuhkan waktu dalam mendapatkan hasil kinerja dosen. Belum lagi standar penilaian yang ada tidak dipahami oleh tim dengan cara yang sama, sehingga menimbulkan inkonsistensi terhadap hasil penilaian. Pelaporan kinerja dosen haruslah sudah diserahkan kepada pimpinan sebelum masuk semester berikutnya, karena laporan tersebut akan menjadi acuan evaluasi bagi pihak rektorat dan lembaga di bawahnya yang terkait dengan tugas dosen. Untuk kemudian nanti hasil evaluasi diumumkan pada rapat dosen pada semester berikutnya dalam bentuk raport kinerja dosen. Pentingnya hal ini untuk diselesaikan tepat waktu menjadi dinamika tersendiri bagi BPMI dalam hal ini lembaga yang melaporkan. Tim BPMI terdiri dari 5 orang, perwakilan fakultas dan vokasi serta 2 orang yang terdiri dari ketua BPMI dan kepala bagian akreditasi.

Dengan melihat dinamika yang terjadi maka perlu dibuat sebuah sistem yang bisa memfasilitasi pengukuran kinerja dosen dan memberikan hasil pengukuran secara cepat dan terintegrasi. dengan sebuah metode yang dapat menyelesaikan penilaian dengan cepat dan mendapatkan hasil yang dapat

dipertanggung jawabkan. Untuk itu metode *Naïve Bayes classifier* dapat menjadi solusi untuk membantu mengatasi permasalahan yang ada. *Naïve Bayes classifier* menurut R. Fieldman dalam (Lestari & Saepudin, 2021) menyatakan merupakan algoritma yang digunakan untuk mencari probabilitas tertinggi untuk mengklasifikasikan data uji pada kategori yang paling tepat. (Zulfikar & Fahmi, 2019) menyatakan bahwa algoritma *Naïve Bayes* mengasumsikan semua atribut adalah *independent* atau bebas yang tidak terpengaruh oleh *variable* kelas. Dengan demikian penggunaan *Naïve Bayes* dapat memudahkan proses penentuan nilai kinerja dosen di Universitas Serang Raya.

Pengukuran kinerja dosen akan memanfaatkan *Sample* data dari dosen yang sudah diukur pada semester sebelumnya, data tersebut akan menjadi *data training* untuk membentuk model probabilitasnya. Setelah model tersebut dibuat diuji akurasi dengan menggunakan *data test*. *Data test* dapat diambil dari *data training* kemudian di bandingkan hasil pengukurannya, yang aktual dan hasil dari model. Dengan adanya *system* perhitungan ini diharapkan dapat mengatasi permasalahan yang ada dan dapat memberikan hasil yang dapat dipertanggung jawabkan secara objektif.

Tujuan yang ingin di capai dalam penelitian antara lain :

1. Membuat rumusan untuk menentukan pola perhitungan kinerja dosen dengan menggunakan algoritma *Naïve Bayes classifier*
2. Membuat algoritma berdasarkan pola pengolahan data dengan menggunakan algoritma *naïve bayes*
3. Menentukan penilain kinerja dosen dengan menggunakan *Naïve Bayes*

II. TINJAUAN PUSTAKA

(Rahutomo et al., n.d.) dengan judul Eksperimen *Naïve Bayes* pada deteksi berita *hoax* berbahasa Indonesia. Website dan blog terkenal sebagai media penayangan berita dalam berbagai bidang seperti penayangan berita. Validitas artikel berita dapat bersifat valid ataupun palsu. Berita palsu disebut juga dengan *hoax news*. Tujuan pembuatan berita *hoax* ini adalah untuk membujuk, memanipulasi, mempengaruhi pembaca berita untuk melakukan hal-hal yang bertentangan atau mencegah tindakan yang benar. Pada penelitian ini mengusulkan untuk melakukan eksperimen klasifikasi *Naïve Bayes* pada deteksi berita *hoax* berbahasa Indonesia. Penelitian ini menggunakan *dataset* sendiri sebanyak 600 berita antara berita *valid* dan *hoax*. Tiga pembaca berita melakukan klasifikasi manual. Sistem yang dibangun dapat mengklasifikasikan berita daring berbahasa Indonesia dengan fitur *term frequency* dan

algoritma klasifikasi *Naïve Bayes* dengan menggunakan komponen *library* PHP-ML atau *PHP-Machine Learning*. Berdasarkan hasil uji coba secara statis, sistem ini menghasilkan akurasi sebesar 82,6% dan pengujian secara dinamis persentase kesesuaian dengan sistem 68,33%. *Dataset* disediakan terbuka sehingga dapat diakses oleh peneliti lainnya dan dapat dijadikan *baseline* pada penelitian-penelitian berikutnya.

(Annur, 2018) meneliti tentang “Klasifikasi Masyarakat Miskin Menggunakan Metode *Naïve Bayes*”. Permasalahan utama dalam upaya pengurangan kemiskinan saat ini terkait dengan adanya fakta bahwa pertumbuhan ekonomi tidak tersebar secara merata. Penelitian akan melakukan klasifikasi berdasarkan data penduduk miskin yang diperoleh dari Kecamatan Tibawa dengan menggunakan teknik data mining. Atribut yang akan digunakan dalam melakukan klasifikasi penduduk adalah Umur, Pendidikan, Pekerjaan, Penghasilan, Tanggungan, Status (Kawin/Belum Kawin). Metode yang akan digunakan adalah metode *Naïve Bayes Classifier*, yang merupakan salah satu Teknik pengklasifikasian dalam data mining. Berdasarkan penelitian yang dilakukan dihasilkan kesimpulan bahwa, Sistem klasifikasi masyarakat miskin di wilayah pemerintahan Kecamatan Tibawa Kab. Gorontalo dapat direkayasa dan Berdasarkan hasil pengujian confusion matrix dengan teknik split validasi, penggunaan metode klasifikasi *Naïve Bayes* terhadap *dataset* yang telah diambil pada objek penelitian diperoleh tingkat akurasi sebesar 73% atau termasuk dalam kategori *Good*. Sementara nilai *Precision* sebesar 92% dan *Recall* sebesar 86%.

(Asfi et al., 2020). Dengan judul “Implementasi *Naïve Bayes Classifier* sebagai Rekomendasi Pembimbing Skripsi” Dosen pembimbing memiliki peran mendampingi mahasiswa dalam proses penyusunan skripsi. Penentuan awal calon pembimbing diberbagai kampus seringkali bersifat subjektif dan kebijakan langsung oleh ketua program studi ataupun dipilih langsung oleh mahasiswa. Penentuan dosen pembimbing terkadang belum dilakukan penyesuaian antara bidang ilmu skripsi yang dipilih mahasiswa dengan kompetensi dosen yang bersangkutan. Oleh karena itu, diperlukan suatu sistem pendukung keputusan dengan penerapan metode yang tepat untuk menentukan dosen pembimbing. Algoritma *Naïve Bayes Classifier* merupakan sebuah metode klasifikasi. Pada penelitian ini digunakan acuan *data training* dosen pembimbing 1 sebanyak 217 dan *data training* dosen pembimbing 2 sebanyak 177, sedangkan data Uji yang digunakan sebanyak 10 data. Kriteria yang digunakan adalah kompetensi, jabatan fungsional dan *homebase* dosen. Implementasi algoritma *Naïve Bayes Classifier* disisipkan dalam

aplikasi sistem pengajuan skripsi terintegrasi SIMASITA CIC. Berdasarkan hasil pengujian algoritma *Naive Bayes Classifier* didapatkan perbandingan tingkat kesesuaian dosen pembimbing 1 sebesar 90% : 10% dan perbandingan tingkat kesesuaian dosen pembimbing 2 sebesar 30% :70 %.

(Senika et al., 2022) dengan judul implementasi *Naive Bayes* dalam penilain kinerja *sales marketing* pada PT. Pachira Distrinusa. Penelitian ini bertujuan untuk membangun sebuah sistem informasi yang dapat menunjang perusahaan dalam pengambilan keputusan khususnya tentang penilaian sales di PT. Pachira Distrinusa. Hal ini dilatar belakangi oleh sulitnya menentukan layak tidaknya seorang sales untuk mendapatkan nilai yang efisien, dikarenakan *system* yang belum terkomputerisasi dan dokumen data karyawan yang bertumpuk. Pada penelitian ini, data yang digunakan adalah data penilaian di PT. Pachira Distrinusa dan metode yang digunakan adalah metode algoritma *Naive Bayes Classifier*. Dan untuk mengetahui seberapa baik algoritma *Naive Bayes Classifier* digunakan pada penelitian ini, maka digunakan perhitungan *RapidMiner* untuk melakukan pengujian. Dari pengujian di *RapidMiner* menghasilkan nilai akurasi sebesar 91,67% dan nilai ROC sebesar 0,979 yang artinya algoritma *Naive Bayes Classifier* sangat baik digunakan pada penelitian ini. Setelah melakukan pengujian menggunakan software *RapidMiner* dan mendapatkan hasil pengujian, kemudian diimplementasikan ke dalam sebuah sistem menggunakan PHP dan MySQL yang dirancang untuk melakukan prediksi penilaian *sales*. Hasil prediksi yang didapatkan dari sistem sudah sesuai dengan hasil perhitungan yang didapatkan dari perhitungan *RapidMiner* dan perhitungan manual. Berdasarkan penelitian yang telah dilakukan bahwa sistem pendukung keputusan yang dibangun ini dapat diaplikasikan pada PT. Pachira Distrinusa sehingga mampu mempermudah dalam menentukan kelayakan penilaian sales di PT. Pachira Distrinusa dengan efisien.

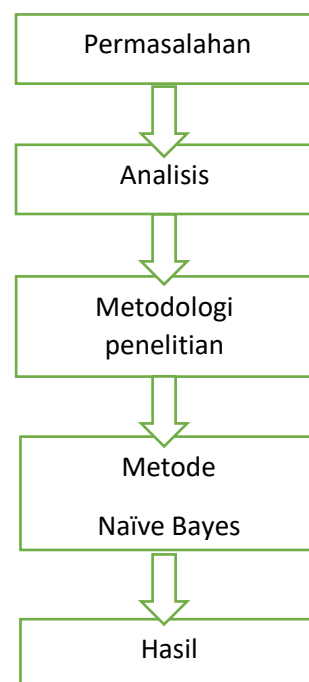
(Sari et al., 2020) meneliti tentang “Analisis dan Penerapan Algoritma *Naive Bayes* untuk Evaluasi Kinerja Karyawan pada PT. Pelita Wira Sejahtera”. Data mining merupakan sebuah teknik dari beberapa bidang ilmu untuk menemukan hubungan yang tidakdiketahui sebelumnya di dalam gudang data sehingga menjadi suatu informasi yang dapat digunakan kemudian. Pada PT. Pelita Wira Sejahtera penilaian terkadang dilakukan secara subjektifdan keterbatasan dalam mengontrol setiap karyawan yang bekerja. Oleh karena itu penulis melakukan analisis data mining pada data-data penilaian karyawan tersebut agar dapat mengetahui mana karyawan yang memiliki kinerja yang sangat baik, baik, cukup, dan kurang. Penulis menggunakan data-data karyawan sebanyak 149

data yang kemudian disajikan kedalam format arff. Dalam melakukan analisis penulis menggunakan alat bantu Tools WEKA. Metode yang di gunakan adalah metode klasifikasi *Naive Bayes* dengan persentasi akurasi terbesar diperoleh dengan menggunakan *Use Training Set Correctly* yaitu sebesar 95.302% , menggunakan *5-Fold Cross Validation Correctly* sebesar 93,9597%, dan menggunakan *10-Fold Cross Validation* sebesar 93.9597%. Sedangkan hasil seleksi atribut menggunakan algoritma *classifier attribute evaluation (Classifier Attribute Eval)* dinyatakan bahwa atribut yang paling berpengaruh terhadap klasifikasi penilaian kineja adalah orientasi pada efisiensi.

III. METODE PENELITIAN

Tahapan Penelitian

Tahapan penelitian memberikan gambaran umum tentang urutan penyelesaian penelitian. Berikut adalah gambaran alur penelitian ini:



Gambar 1. Tahapan penelitian

Jadwal Penelitian

Waktu yang dibutuhkan penelitian ini kurang lebih selama 1 tahun, dimulai dari bulan Maret 2022 sampai dengan bulan Februari 2023.

Naive bayes

Naive Bayes adalah *classifier* sederhana yang didasarkan pada *teorema bayes*. Kelebihan dari Teknik ini adalah mampu mengklasifikasikan dokumen dengan tepat serta mudah dalam mengimpilemntasikannya, Thabta (2009). Olson dan Dursun menjelaskan *Naive Bayes* untuk setiap

kelas keputusan, menghitung probabilitas dengan syarat bahwa semua kelas keputusan benar, mengiangat *vector* informasi. Algoritma *Naïve Bayes* ini mengasumsikan bahwa atribut objek adalah independent. Probabilitas yang terlibat dalam memproduksi perkiraan akhir dihitung sebagai jumlah frekuensi dari “*master*” tabel keputusan (Olson & Delen Dursun, 2008). Dalam bukunya han J. menyatakan “*Bayessian Classffiers* mempunyai tingkat kecepatan dan akurasi yang tinggi Ketika diaplikasikan dalam *database* yang besar” (Han Jiawei et al., 2021).

Algoritma *Naïve Bayes classffier* merupakan algoritma yang digunakan mencari nilai probabilitas tertinggi untuk mengklasifikasikan data uji pada kategori yang paling tepat (Feldman et al., 2007). Persamaan *Naïve Bayes* yang digunakan adalah sebagai berikut:

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

Dimana:

P(H|X) merupakan H di dalam X, atau banyaknya H didalam X

P(X|H) probabilitas X didalam H

P(H) adalah probablitas Prior dari H, dan

P(X) adalah probablitas prior dari X

Analisis Kebutuhan Data

Data menjadi elemen yang menjadi dasar dalam penelitian ini, mendapatkan data kinerja dosen selama satu semester sebagai gambaran untuk diolah dengan menggunakan *Naïve Bayes*. Data yang dikumpulkan adalah data satu semester dikarenakan evaluasi kinerja dosen dilakukan setiap semester sekali di Universitas Serang Raya. Data didapatkan dari biro kepegawaian Universitas Serang Raya sedangkan *instrument* pengukuran yang sedang berjalan diambil dari Badan Penjamin Mutu Internal (BPMI). Dari data yang didapat akan dilakukan perhitungan untuk menemukan pola dengan metode *Naïve Bayes*.

Data yang digunakan pada penelitian diatas hanya data dosen dalam satu fakultas yakni Fakultas Teknologi Informasi. Pada semester ganjil 2020/2022 jumlah data dosen aktif adalah sekitar 34 orang. 34 orang ini adalah dosen dengan status dosen tetap yayasan, sedangkan dosen luar biasa tidak di masukan ke dalam pengukuran. Data yang akan diolah adalah data kinerja dosen di Fakultas Teknologi Informasi selama periode semester ganjil 2021/2022 atau dalam satuan waktu periode tersebut di mulai dari bulan Oktober tahun 2021 sampai dengan bulan Maret 2022. Data tersebut akan menjadi *data set* atau *data training* yang akan di olah dengan *navie bayes classifier*.

IV. HASIL DAN PEMBAHASAN

Hasil dari penelitian ini adalah sebuah *system* perhitungan yang dapat merekomendasikan pemenuhan standar kinerja dosen dalam skala yang sudah ditentukan. Skala pengukuran mengacu kepada bukti kinerja dosen dalam pemenuhan kewajiban tridharma perguruan tinggi. Berikut adalah gambaran aplikasi yang sudah diselesaikan sebagai produk penelitian

Dasar dari perhitungan kinerja adalah aturan yang dibuat oleh badan penjamin mutu internal yang mengacu kepada PO BKD Ildikti dengan beberapa penyesuaian. Namun dalam penelitian ini *indicator* pengukuran kinerja dibatasi hanya beberapa saja yang diambil. Berikut adalah poin-poin yang di ambil dalam aturan kinerja dosen pada penelitian ini :

Tabel 1. Indikator Kinerja

Bidang	Kode atribut	Kegiatan
Pengajaran	-	Jumlah presensi tatap muka <i>real</i>
Penelitian	B0	Tidak melakukan penelitian
	B1	Melakukan penelitian dengan bukti laporan
	B2	Publikasi jurnal/prosding internasional
	B3	Publikasi jurnal/prosiding nasional
	B4	HKI
	B5	Paten
Pengabdian	C0	Tidak melakukan pengabdian
	C1	Melakukan pengabdian dengan bukti laporan
	C2	Publikasi jurnal pengabdian
Penunjang	D0	Tidak ada Penunjang
	D1	Pertemuan ilmiah
	D2	Kepanitiaan internal kampus
	D3	Kepanitian badan eksternal kampus

Berdasarkan indikator kinerja yang sudah diisi dan dibuktikan dengan mengumpulkan berkas-berkas yang dibutuhkan, selanjutnya adalah proses penilaian berdasarkan kinerja tiap-tiap dosen. Pengukuran kinerja di lakukan secara manual oleh staff kepegawaian. Hasil pengukuran di buat dalam skala yang ditunjukan oleh tabel berikut:

Tabel 2. Skala kinerja Dosen

No	Skala
1	Baik Sekali
2	Baik
3	Cukup
4	Kurang

Skala yang ditunjukkan pada tabel 4.2 akan menjadi kelas dalam klasifikasi dan Hasil pengukuran yang di lakukan admin kepegawaian akan menjadi landasan awal perhitungan dengan *Naïve Bayes*, yang kemudian akan di sebut sebagai *dataset* . *dataset* ini akan dijadikan rujukan untuk menghitung probablitas nilai kinerja dosen di periode penilaian berikutnya.

a. *Dataset*

Dataset yang digunakan pada simulasi perhitungan adalah data kinerja dosen Fakultas Teknologi Informasi pada periode pengukuran tahun ajaran 2021-2022 dan pada semester ganjil. Data kinerja dosen tersebut diukur dengan kegiatan tridharma yang terdiri dari pengajaran dalam satuan nilai rata-rata presensi tatap muka *real*, penelitian dengan adanya bukti penelitian begitupun dengan pengabdian dan penunjang, sedangkan label atau kelas untuk data tersebut terdiri dari baik sekali, baik, cukup dan kurang. *Dataset* juga nanti dapat disebut sebagai *data training*, Berikut adalah *data training* pada periode tersebut :

Tabel 3 *Dataset* Dosen FTI Semester Ganjil 2021/2022

No	Nama	A	B	C	D	Nilai
1	AGI	12	B1	C1	D1	Baik
2	AGS	12	-	C1	D1	Kurang
3	AP	13	B1	C1	D1	Baik
4	SUL	15	B1	C2	D1	Baik Sekali
5	TN	14	-	C1	D1	Kurang
6	DS	16	B3	C2	D1	Baik Sekali
7	ZM	14	B4	C1	D1	Baik
8	SIS	15	B3	C1	D1	Baik
9	MS	14	B1	C1	D1	Cukup
10	HTS	13	B3	C2	D1	Baik Sekali
11	AKS	14	B3	C1	D1	Baik
12	SUM	14	B3	C2	D1	Baik Sekali
13	NGT	14	0	C1	D1	Cukup
14	SS	16	0	0	0	Kurang
15	DB	16	0	C1	D1	Cukup
16	VR	14	B3	C2	D1	Baik Sekali
17	HS	14	B2	C2	D1	Baik Sekali
18	SAE	16	B1	C1	D1	Baik
19	ANH	15	B3	C1	D1	Baik Sekali
20	SUT	14	B1	C2	D1	Baik
21	SD	16	0	C1	D1	Cukup
22	DF	12	B3	C1	D1	Baik
23	RN	14	B4	C1	D1	Baik
24	MI	15	B1	C1	D1	Baik
25	SUH	13	B3	C1	D2	Baik
26	SW	14	B1	C1	D1	Baik
27	DK	16	B3	C2	D1	Baik Sekali

No	Nama	A	B	C	D	Nilai
28	ES	15	B1	C1	D1	Baik
29	HG	12	0	0	0	Kurang
30	ADJ	10	B1	C1	D1	Cukup
31	AK	12	B1	C1	D1	Baik
32	EP	16	B1	C1	D1	Baik
33	AS	14	0	0	D1	Kurang
34	SD	15	B3	C2	D1	Baik Sekali

Keterangan :

- Smt : Semester
- A : Bidang Pengajaran
- B : Bidang Penelitian
- C : Bidang Pengabdian kepada Masyarakat
- D : Bidang penunjang

b. Perhitungan

1. Menghitung probabalitas kelas
Untuk menghitung probabilitas kelas maka dihitung jumlah frekuensi kelas muncul dibagi dengan total *dataset* . Adapun perhitungannya sebagai berikut:

$$P(Y = \text{baik sekali}) = \frac{9}{34} = 0,26470$$

$$P(Y = \text{baik}) = \frac{15}{34} = 0,441176$$

$$P(Y = \text{cukup}) = \frac{5}{34} = 0,147058$$

$$P(Y = \text{kurang}) = \frac{5}{34} = 0,147058$$

Sehingga didapatkan hasil sebagai berikut:

Tabel 4. probabilitas kelas

Nama kelas	Nilai P(Y)
Baik sekali	0,26470
Baik	0,441176
Cukup	0,147058
Kurang	0,147058

2. Probabilitas atribut

Langkah selanjutnya adalah mencari nilai probabilitas dari setiap atribut, mengacu kepada *data training* yang ada. Perhitungannya dengan membagi nilai frekuensi atribut dengan banyaknya data. adapun perhitungannya sebagai berikut: Khusus untuk atribut pengajaran (A) karena nilainya numerik (*continue*), maka perhitungan probabilitas menggunakan distribusi *Gaussian*. Adapun rumus nya adalah sebagai berikut:

$$P(h_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(h_j - \mu)^2}{2\sigma^2}$$

Keterangan:

- P(hj) : Probabilitas atribut yang

σ : Standar Deviasi
μ : Mean

maka untuk menghitung Probabilitas atribut pengajaran maka harus hitung terlebih dahulu mean (rata-rata) dan nilai standar deviasi nya. Hasil dari mencari standar deviasi dan mean di tunjukan pada tabel berikut :

Tabel 5. Mean dan Standar Deviasi Pengajaran

Kelas	Mean	Standar deviasi
Baik sekali	14,6667	1
Baik	13,9333	1,3345
Cukup	14	2,44949
Kurang	13,6	1,67332

Dari tabel di atas didapatkan nilai mean dan nilai standar deviasinya, nilai tersebut nanti digunakan untuk menghitung distribusi *Gaussian data test*. Karena nilai numerik dari data uji harus dimasukan dalam rumus distribusi gaussian tersebut. Maka selanjutnya perhtiungan untuk atribut Penelitian. Kerena datanya berupa kategorial (*polynomial*) maka perhitunganya dilakukan dengan membagi jumlah atribut yang muncul pada tiap kelas. Untuk contoh peritunganya yang diambil adalah nilai atribut B1 yang artinya melakukan penelitian dengan bukti laporan, maka perhitunganya adalah sebagai berikut:

$$P(\text{penelitian} = B1 | Y = \text{baik sekali}) = \frac{1}{9} = 0,11111$$

Penjelasan perhitungan di atas adalah untuk mendapatkan peluang B1 dalam kelas baik sekali dari data training di atas, nilai 1 adalah banyaknya nilai atribut B1 didalam kelas Baik sekali, sedangkan nilai 9 adalah jumlah seluruh kelas baik sekali. Perhitungan diteruskan hanya untuk nilai atribut B1 di dalam kelas lainnya:

$$P(\text{penelitian} = B1 | Y = \text{baik}) = \frac{9}{15} = 0,6$$

$$P(\text{penelitian} = B1 | Y = \text{cukup}) = \frac{2}{5} = 0,4$$

$$P(\text{penelitian} = B1 | Y = \text{kurang}) = \frac{0}{5} = 0$$

Berdasarkan cara perhitungan di atas maka dilakukan perhitungan untuk semua nilai atribut bidang B (penelitian), sehingga didapatkan hasil yang ditunjukan tabel sebagai berikut:

Tabel 6 nilai probablitas atribut bidang Penelitian

Atribut B	BS	B	C	K
B0	0	0	0,6	1
B1	0,111111	0,6	0,4	0
B2	0,111111	0	0	0
B3	0,777778	0,266667	0	0
B4	0	0,133333	0	0
B5	0	0	0	0

Pada tabel probabilitas atribut terlihat bahwa masih banyak kolom yang bernilai 0, dalam *Naïve Bayes* hal ini menjadi anomali karena tidak boleh terdapat nilai 0 dalam probabilitas karena nilai 0 menyatakan “tidak mungkin terjadi”. *Anomaly* tersebut kemudian dapat di koreksi dengan menggunakan Teknik perhitungan *Laplacian correction*. *Laplacian correction* atau *Laplacian smoothie* adalah Teknik perhtitungan untuk menyiasati supaya probabilitas pada perhitungan *navie bayes classifier* tidak bernilai 0, dikarenakan tidak adanya data untuk kategori tertentu dalam sebuah *class* (Rizki et al., 2021). Persamaan *Laplacian correction* adalah sebagai berikut :

$$P(c, t) = \frac{1 + N(ct, D)}{N(c, D) + |V|}$$

Dimana:

P(c,t) : Probabilitas nilai yang akan ditemukan

N(ct, D) : jumlah nilai atribut

N(c,D) : jumlah sampel

|V| : jumlah sample pada atribut P

Maka untuk menghilangkan nilai 0 pada tabel nilai probabilitas atribut penelitian (B) di atas dilakukan perhitungan lagi menggunakan Teknik *Laplacian correction*. Untuk menggunakan *Laplacian correction* yang perlu diperhatikan adalah kelompok nilai atribut yang bernilai 0 berada pada kelas yang mana, apabila dalam kelas tersebut nilai probabilitas atributnya tidak ada yang bernilai 0 maka tidak perlu melakukan *Laplacian correction* pada nilai atribut di dalam kelas tersebut, namun apabila sebaliknya maka dilakukan perhitungan *Laplacian correction* untuk setiap nilai atribut yang ada pada kelas tersebut. Sebagai contoh tabel nilai probabilitas atribut penelitian di atas, semua kelas memiliki nilai atribut 0, sehingga *Laplacian correction* akan dilakukan pada semua atribut pada setiap kelas. Untuk perhitungannya adalah sebagai berikut:

Pada tabel di atas pada kelas Baik sekali (BS) ada 6 nilai kriteria yaitu B1 = 0,11111, B2 = 0,11111; B3 = 0,77778; B4 = 0 dan B5 = 0. Karena ada nilai 0 pada kelas Baik sekali (BS) maka *Laplacian correction* dapat di terapkan pada kelas ini. Untuk menerapkan *Laplacian correction* maka posisi probabilitasnya dihitung sebelum mendapatkan nilai desimalnya misalkan untuk B1 = 0,11111 di kembalikan ke persamaan sebelum desimalnya yaitu jumlah B1 dalam Baik sekali dibagi dengan total baik sekali menjadi 1/9, setelah itu nilai 1/9 masukan ke dalam persamaan *Laplacian correction*, sehingga menjadi :

$$\begin{aligned}
 P(\text{penelitian} = B1 | \text{Baik sekali}) &= \frac{1 + 1}{9 + 6} = \frac{2}{15} \\
 &= 0,13333
 \end{aligned}$$

Kemudian perhitungan *laplacian correction* di atas diterapkan ke semua nilai atribut dalam kelas Baik Sekali (BS) yang di dalamnya nanti ada nilai probabilitas 0 :

$$\begin{aligned}
 P(\text{penelitian} = B2 | \text{Baik sekali}) &= \frac{1 + 1}{9 + 6} = \frac{2}{15} \\
 &= 0,13333
 \end{aligned}$$

$$\begin{aligned}
 P(\text{penelitian} = B3 | \text{Baik sekali}) &= \frac{1 + 7}{9 + 6} = \frac{8}{15} \\
 &= 0,53333
 \end{aligned}$$

Untuk B4, B5 dan B6 karena tidak terpilih nilai atributnya, maka frekuensi nilainya adalah 0 sehingga berapapun pembagiannya hasil akhirnya tetap bernilai 0. B4, B5 dan B6 adalah alas an atribut pada kelas ini di masukan ke dalam persamaan *Laplacian correction* karena 0 dalam probabilitas ada tidak mungkin, sehingga kasus seperti ini dapat ditemukan solusinya dengan *Laplacian correction*. Dan hasil perhitungannya adalah sebagai berikut:

$$\begin{aligned}
 P(\text{penelitian} = B4 | \text{Baik sekali}) &= \frac{1 + 0}{9 + 6} = \frac{1}{15} \\
 &= 0,06667
 \end{aligned}$$

$$\begin{aligned}
 P(\text{penelitian} = B5 | \text{Baik sekali}) &= \frac{1 + 0}{9 + 6} = \frac{1}{15} \\
 &= 0,06667
 \end{aligned}$$

$$\begin{aligned}
 P(\text{penelitian} = B6 | \text{Baik sekali}) &= \frac{1 + 0}{9 + 6} = \frac{1}{15} \\
 &= 0,06667
 \end{aligned}$$

Setelah dilakukan persamaan *Laplacian correction* untuk mengatasi nilai 0 dalam probabilitas atribut penelitian maka didapatkan nilai probabilitas baru pada kelas baik sekali (BS) B0 = 0,06667; B1= 0,13333; B2 = 0,13333; B3 = 0,53333; B4 yang tadinya bernilai 0 menjadi 0,06667 dan B5 yang juga 0 menjadi 0,06667. *Laplacian correction* di semua kelas pada atribut penelitian karena setiap kelas memiliki nilai probabilitas atribut 0. Setelah dilakukan *Laplacian correction* pada semua kelas pada atribut penelitian di dapatkan hasil seperti pada tabel berikut :

Tabel 7. Nilai probabilitas penelitian dengan *Laplacian correction*

Atribut B	Baik sekali	Baik	Cukup	Kurang
B0	0,0666 6667	0,0476 19048	0,3636 36364	0,5454 54545
B1	0,1333 3333	0,4761 90476	0,2727 27273	0,0909 09091
B2	0,1333 3333	0,0476 19048	0,0909 09091	0,0909 09091
B3	0,5333 3333	0,2380 95238	0,0909 09091	0,0909 09091
B4	0,0666 6667	0,1428 57143	0,0909 09091	0,0909 09091
B5	0,0666 6667	0,0476 19048	0,0909 09091	0,0909 09091

Langkah selanjutnya adalah menghitung nilai atribut bidang C (pengabdian kepada masyarakat). Karena sifatnya atributnya deksrit atau *polynomial* maka cara perhitungannya serupa dengan cara peritugan probabilitas bidang B, sehingga hasil yang didapatkan adalah sebagai berikut:

Tabel 8 Nilai probabilitas atribut Pengabdian

atribut C	BS	B	C	K
C0	0	0	0	0,6
C1	0,111111	0,933333	1	0,4
C2	0,888889	0,066667	0	0

Setelah dicari probabilitas nya, pada tabel di atas menunjukkan masih ada probabilitas atribut dalam kelas yang bernilai 0. Maka di lakukan *Laplacian correction* juga terhadap kelas yang mamiliki nilai probabilitas atribut 0 tersebut. Sehingga hasilnya sebagai berikut:

Tabel 9. Nilai probabilitas pengabdian dengan *Laplacian correction*

atribut C	BS	B	C	K
C0	0,083333 33	0,055555 56	0,12 5	0,5
C1	0,166666 67	0,833333 33	0,75	0,37 5
C2	0,75	0,111111 11	0,12 5	0,12 5

Untuk perhitungan probabilitas atribut terakhir adalah bidang D yaitu penunjang. Nilai atribut penunjang juga binominal juga dihitung dengan cara yang sama dan hasil nya adalah sebagai berikut;

Tabel 10 Nilai probabilitas atribut bidang Penunjang

atribut D	BS	B	C	K
D0	0	0	0	0,4
D1	1	0,933333	1	0,6
D2	0	0,066667	0	0

Tabel probabilitas atribut penunjang pun masih mempunyai nilai 0 pada probabilitas nilai atributnya. Maka pada tabel di atas pun dilakukan *Laplacian correction* untuk menghilangkan nilai 0 pada kelas yang memiliki probabilitas atribut 0, sehingga menjadi seperti berikut:

Tabel 11 Probabilitas atribut penunjang dengan *Laplacian correction*

atribut D	BS	B	C	K
D0	0,08333 333	0,05555 556	0,1 25	0,3 75
D1	0,83333 333	0,83333 333	0,7 5	0,5
D2	0,08333 333	0,11111 111	0,1 25	0,1 25

3. Menghitung *likelihood* / kemungkinan

Langkah selanjutnya melakukan eksperimen perhitungan dengan data uji yang di ambil dari data training, untuk mensimulasikan cara mencari *likelihood*. Untuk menghitung *likelihood* dengan melakukan perkalian probabilitas nilai atribut dalam kelas kemudian dikalikan lagi dengan probabilitas kelas. Untuk melakukan simulasi maka ambil satu data dari data training misalkan data yang akan jadi data uji adalah sebagai berikut:

Tabel 12 *Data test*

Nama	Prodi	A	B	C	D	nilai
ANH	SK	15	B3	C1	D1	?

Langkah pertama adalah mencari probabilitas atribut yang bertipe numerik yaitu pengajaran (A). untuk mencari probabilitasnya maka dilakukan dengan persamaan distribusi *gaussian*:

$$P(h_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(h_j - \mu)^2}{2\sigma^2}$$

Sehingga untuk alternatif ANH yang nilai atribut A nya adalah 15 dimasukan ke persamaan distribusi *gaussian* menjadi:

$$P(A = 15 | Baik sekali) = \frac{1}{\sqrt{2 \times 3.14 \times 1^2}} \exp - \frac{(15 - 14.667)^2}{2 \times 1^2} = 0,3775$$

$$P(A = 15 | Baik) = \frac{1}{\sqrt{2 \times 3.14 \times 1,3345^2}} \exp - \frac{(15 - 13.933)^2}{2 \times 1,3345^2} = 0,251$$

$$P(A = 15 | cukup) = \frac{1}{\sqrt{2 \times 3.14 \times 2,4495^2}} \exp - \frac{(15 - 14)^2}{2 \times 2,4495^2} = 0,2346$$

$$P(A = 15 | Kurang) = \frac{1}{\sqrt{2 \times 3.14 \times 1,6733^2}} \exp - \frac{(15 - 13,6)^2}{2 \times 1,6733^2} = 0,2174$$

Dari perhitungan di atas ditemukan nilai P(A=15| Kelas), maka Langkah selanjutnya adalah menghitung *likelihood* dengan persamaan :

Likelihood (BS)= P(A=15| Baik sekali) x P(B=B3|baik sekali) x P(C= C1 | baik sekali) x P(D = D1 | Baik sekali) x P(Baik sekali). Sehingga dapat dihitung sebagai berikut:

$$\begin{aligned} \text{likelihood}(\text{baik sekali}) &= 0,3775 \times 0,53333 \times 0,166667 \times 0,83333 \times 0,26470 \\ &= 0,0074 \end{aligned}$$

Setelah itu lalu di hitung *likelihood* untuk semua kelas lainnya :

$$\begin{aligned} \text{likelihood}(\text{baik}) &= 0,251 \times 0,2380 \times 0,83333 \times 0,83333 \times 0,4412 \\ &= 0,01831 \end{aligned}$$

$$\begin{aligned} \text{likelihood}(\text{cukup}) &= 0,2346 \times 0,0909091 \times 0,75 \times 0,75 \times 0,1471 \\ &= 0,00176 \end{aligned}$$

$$\begin{aligned} \text{likelihood}(\text{kurang}) &= 0,2174 \times 0,0909091 \times 0,375 \times 0,5 \times 0,1471 \\ &= 0,000545 \end{aligned}$$

4. Melakukan normalisasi *likelihood* “baik sekali”

$$\begin{aligned} \text{baik sekali} &= \frac{0,0074}{0,0074 + 0,01831 + 0,00176 + 0,000545} \\ &= \frac{0,0074}{0,028015} = 0,2641 \end{aligned}$$

Probabilitas untuk “baik”

$$\begin{aligned} \text{baik} &= \frac{0,01831}{0,0074 + 0,01831 + 0,00176 + 0,000545} \\ &= \frac{0,01831}{0,028015} = 0,6535 \end{aligned}$$

Probabilitas untuk “cukup”

$$\begin{aligned}
 & \text{cukup} \\
 & = \frac{0,00176}{0,0074 + 0,01831 + 0,00176 + 0,000545} \\
 & = \frac{0,00176}{0,028015} = 0,0628
 \end{aligned}$$

Probabilitas untuk “kurang”

$$\begin{aligned}
 & \text{baik} \\
 & = \frac{0,000545}{0,0074 + 0,01831 + 0,00176 + 0,000545} \\
 & = \frac{0,000545}{0,028015} = 0,0194
 \end{aligned}$$

Maka dapat ditarik kesimpulan dari data test dengan nama dosen anharudin mendapatkan nilai lebih besar untuk kelas baik dengan nilai sebesar 0,6406 atau dalam persentasi menjadi 64,06%. Untuk baik sekali sebesar 0.2589 atau 25,89 %, sedangkan cukup 0,0433 atau 4% dan kurang 0,057 atau 5 %. Namun hasil yang didapat tidak sesuai dengan data yang ada pada data training sehingga perlu melakukan pengujian akurasi terhadap perhitungan *Naïve Bayes* classifier yang sudah di lakukan.

Pengujian akurasi

Untuk membuktikan validitas *Naïve Bayes classifier* di atas maka dilakukan pengujian akurasi. Dalam pengujian akurasi dapat menggunakan beberapa metode pengujian misalkan dengan metode *confusion matrix*.

Cara pengujian akurasi dilakukan dengan menghitung jumlah prediksi yang benar di bagi dengan jumlah data dikalikan dengan 100%. Maka penyelesaian untuk akurasi dapat di cari dengan menggunakan persamaan berikut :

$$\text{akurasi} = \frac{\text{jumlah prediksi benar}}{\text{jumlah dataset}} \times 100\%$$

Untuk memulai menghitung akurasi terlebih dahulu menghitung nilai prediksi seluruh *dataset* kemudian membandingkan dengan data set yang aktual/kenyataannya. Maka perbandingan ini di tunjukan pada tabel berikut:

Tabel 13. Nilai Likelihood Berdasarkan kelas

No	Nama	Baik Sekali	Baik	Cukup	Kurang
1	AGI	5,58768E-05	0,017646	0,004122	0,00049
2	AGS	2,79384E-05	0,001765	0,005495	0,002937
3	AP	0,000487757	0,039462	0,005752	0,000751
4	SUL	0,008326741	0,004882	0,000882	0,000182
5	TN	0,000783159	0,005033	0,007669	0,004509
6	DS	0,014475147	0,001013	0,000229	9,22E-05
7	ZM	0,000783159	0,0151	0,001917	0,000751
8	SIS	0,007401548	0,018308	0,001764	0,000545
9	MS	0,001566319	0,050332	0,005752	0,000751

No	Nama	Baik Sekali	Baik	Cukup	Kurang
10	HTS	0,00877962	0,002631	0,00032	0,00025
11	AKS	0,006265275	0,025166	0,001917	0,000751
12	SUM	0,028193736	0,003355	0,00032	0,00025
13	NGT	0,000783159	0,005033	0,007669	0,004509
14	SS	2,01044E-05	6,75E-06	0,000153	0,001659
15	DB	0,000402087	0,001519	0,005495	0,001659
16	VR	0,028193736	0,003355	0,00032	0,00025
17	HS	0,007048434	0,000671	0,00032	0,00025
18	SAE	0,000804175	0,015192	0,004122	0,000276
19	ANH	0,007401548	0,018308	0,001764	0,000545
20	SUT	0,007048434	0,006711	0,000959	0,00025
21	SD	0,000402087	0,001519	0,005495	0,001659
22	DF	0,000223507	0,008823	0,001374	0,00049
23	RN	0,000783159	0,0151	0,001917	0,000751
24	MI	0,001850387	0,036615	0,005292	0,000545
25	SUH	0,000195103	0,002631	0,00032	0,000188
26	SW	0,001566319	0,050332	0,005752	0,000751
27	DK	0,014475147	0,001013	0,000229	9,22E-05
28	ES	0,001850387	0,036615	0,005292	0,000545
29	HG	1,39692E-06	7,84E-06	0,000153	0,002937
30	ADJ	3,65095E-08	0,000655	0,001516	7,64E-05
31	AK	5,58768E-05	0,017646	0,004122	0,00049
32	EP	0,000804175	0,015192	0,004122	0,000276
33	AS	0,00039158	0,000336	0,001278	0,006012
34	SD	0,033306964	0,002441	0,000294	0,000182

Nilai yang diambil adalah nilai kemungkinan dari kelas yang terbesar nilainya. Maka itu menjadi kelas penilaian bagi dosen, untuk mengetahui nilai prediksi dan membandingkannya dengan nilai aktual, maka dibuat sebuah tabel sebagai berikut:

Tabel 14. Perbandingan hasil aktual dan prediksi

No	Nama	Aktual	Prediksi
1	AGI	Baik	Baik
2	AGS	Kurang	Cukup
3	AP	Baik	Baik
4	SUL	Baik Sekali	Baik Sekali
5	TN	Kurang	Cukup
6	DS	Baik Sekali	Baik Sekali
7	ZM	Baik	Baik
8	SIS	Baik	Baik
9	MS	Cukup	Baik
10	HTS	Baik Sekali	Baik Sekali
11	AKS	Baik	Baik
12	SUM	Baik Sekali	Baik Sekali
13	NGT	Cukup	Cukup
14	SS	Kurang	Kurang

15	DB	Cukup	Cukup
16	VR	Baik Sekali	Baik Sekali
17	HS	Baik Sekali	Baik Sekali
18	SAE	Baik	Baik
19	ANH	Baik	Baik
20	SUT	Baik Sekali	Baik Sekali
21	SD	Cukup	Cukup
22	DF	Baik	Baik
23	RN	Baik	Baik
24	MI	Baik	Baik
25	SUH	Baik	Baik
26	SW	Baik	Baik
27	DK	Baik Sekali	Baik Sekali
28	ES	Baik	Baik
29	HG	Kurang	Kurang
30	ADJ	Cukup	Cukup
31	AK	Baik	Baik
32	EP	Baik	Baik
33	AS	Kurang	Kurang
34	SD	Baik Sekali	Baik Sekali

Dari 34 data uji yang juga merupakan *dataset* didapatkan hasil yang cocok antara nilai aktual dan nilai prediksi berjumlah 31 dan yang tidak sesuai adalah 3. Maka untuk mengukur akurasi berdasarkan persamaan di atas dilakukan perhitungan sebagai berikut:

$$\begin{aligned} \text{akurasi} &= \frac{31}{34} \times 100\% = 0.9117 \times 100\% \\ &= 91,17\% \end{aligned}$$

Sedangkan nilai error pada diukur dengan cara yang sama menggunakan jumlah ketidakcocokan antara nilai actual dan nilai prediksi. Dapat dihitung sebagai berikut :

$$\text{error} = \frac{3}{34} \times 100\% = 0.088 \times 100\% = 8.8\%$$

Nilai error yang terjadi akibat ketidakcocokan adalah sebesar 8,8 %. Dengan melihat nilai akurasi yang sudah dihitung di atas maka hasil dari implementasi *Naïve Bayes classifier* ini sudah sesuai dengan harapan.

V. PENUTUP

Kesimpulan

Kesimpulan yang dapat di ambil dari penelitian ini antara lain:

1. Untuk mengukur kinerja dosen dalam satu semester dengan menggunakan algoritma *Naïve Bayes classifier* mengolah *dataset* sebanyak 34 data dari dosen di Fakultas

Teknologi Informasi yang aktif. Atribut kinerja dibuat berdasarkan unsur tridharma yang terdiri dari pengajaran, penelitian dan pengabdian. Pembatasan dilakukan dalam membuat nilai atribut. Tipe yang digunakan adalah tipe numerik atau kotinyu dan nilai diskrit atau kategorial dengan kelas yang terdiri dari baik sekali, baik, cukup dan kurang. Perhitungan disesuaikan dengan konsep *naïve bayes*, mulai dari mencari probabilitas kelas, probabilitas atribut dan kemudian mencari nilai *likelihood* atau kemungkin, pada tahap selanjutnya hasilnya dapat dinormalisasi ke dalam pembagian nilai prediksi dalam bentuk persentasi.

2. Dari hasil pengujian yang dilakukan tingkat akurasi dari pengujian data uji yang juga merupakan *dataset* menghasilkan nilai akurasi sebesar 91,17 %. Tingkat akurasi ini sudah sesuai dengan yang diharapkan.

Saran

Adapun saran yang dapat diberikan antara lain:

1. Penelitian ini merupakan awal dari rangkaian penelitian kami tentang data mining klasifikasi, sehingga kami dapat membumat penelitian lanjutan terkait tema serupa.
2. Penelitian ini dapat menjadi pemicu untuk peneliti-peneliti lain dapat melanjutkan atau mengembangkan konsep dan mungkin menjajdikan kearah pembuatan produk seperti *machine learning*.

DAFTAR PUSTAKA

- Annur, H. (2018). Klasifikasi Masyarakat Miskin Menggunakan Metode Naïve Bayes. In *Agustus* (Vol. 10, Issue 2).
- Asfi, M., Fitrianiingsih, N., Pembimbing, D., Skripsi, N., & Bayes, C. (2020). *Implementasi Algoritma Naive Bayes Classifier sebagai Sistem Rekomendasi Pembimbing Skripsi*. 5(1).
<https://doi.org/10.30743/infotekjar.v5i1.2536>
- Feldman, R., Ronen, Sanger, & James. (2007). *The text mining handbook: Advanced approaches in analyzing unstructured data*.
- Han Jiawei, pei Jian, & Tong Hanghang. (2021). *Data mining: concept and techniques* (2nd ed.). Kaufmann.
- Lestari, S., & Saepudin, S. (2021). *Analisis Sentimen Vaksin Sinovac Pada Twitter Menggunakan Algoritma Naive Bayes*.
<https://vaksin.kemkes.go.id/>
- Olson, D. L., & Delen Dursun. (2008). *Advanced Data Mining Techniques*. Springer Berlin

- Heidelberg. <https://doi.org/10.1007/978-3-540-76917-0>
- Rahutomo, F., Yanuar Risca Pratiwi, I., Mayangsari Ramadhani, D., & Negeri Malang Jalan Soekarno Hatta No, P. (n.d.). *Eksperimen Naïve Bayes pada Deteksi Berita Hoax Berbahasa Indonesia Naïve Bayes's Experiment on Hoax News Detection in Indonesian Language*.
- Rizki, M., Arhami, M., Huzeni, D., Teknologi, J., Dan, I., Politeknik, K., Lhokseumawe, N., Aceh, J. B., & Km, M. (2021). Perbaikan Algoritma *Naive Bayes Classifier* Menggunakan Teknik Laplacian Correction. In *Jurnal Teknologi* (Vol. 21, Issue 1).
- Sari, V. N., Astri, L. Y., & Rasywir, E. (2020). Analisis dan Penerapan Algoritma Naive Bayes untuk Evaluasi Kinerja Karyawan pada PT. Pelita Wira Sejahtera. *Jurnal Ilmiah Mahasiswa Teknik Informatika*, 2(1), 53–68.
- Senika, A., Rasiban, R., & Iskandar, D. (2022). Implementasi Metode *Naïve Bayes* Dalam Penilaian Kinerja Sales Marketing Pada PT. Pachira Distrinusa. *Jurnal Media Informatika Budidarma*, 6(1), 701. <https://doi.org/10.30865/mib.v6i1.3331>
- Zulfikar, M., & Fahmi, H. (2019). Penerapan Sistem Pendukung Keputusan Dengan Metode *Naïve Bayes* Dalam Menentukan Kualitas Bibit Padi Unggul Pada Balai Pertanian Pasar Miring. *Jurnal Nasional Komputasi Dan Teknologi Informasi*, 2(2).