

**IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK
KLASIFIKASI BUKU PERPUSTAKAAN PADA SISTEM
INFORMASI PERPUSTAKAAN KABUPATEN BANTUL**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

M ZUFAR TAQUIDDIN

17.11.1304

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2022

**IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK
KLASIFIKASI BUKU PERPUSTAKAAN PADA SISTEM
INFORMASI PERPUSTAKAAN KABUPATEN BANTUL**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
M ZUFAR TAQIUDDIN
17.11.1304

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2022**

HALAMAN PERSETUJUAN

SKRIPSI

**IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK KLASIFIKASI
BUKU PERPUSTAKAAN PADA SISTEM INFORMASI PERPUSTAKAAN
KABUPATEN BANTUL**

yang disusun dan diajukan oleh

M Zufar Taqluddin

17.11.1304

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 25 November 2022

Dosen Pembimbing,



Anna Balta, M. Kom

NIK. 190302290

HALAMAN PENGESAHAN

SKRIPSI

**IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK KLASIFIKASI
BUKU PERPUSTAKAAN PADA SISTEM INFORMASI PERPUSTAKAAN
KABUPATEN BANTUL.**

yang disusun dan diajukan oleh

M Zufar Taqiuddin

17.11.1304

Telah dipertahankan di depan Dewan Penguji
pada tanggal 25 November 2022

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dina Maulna, M. Kom
NIK. 190302250

Anna Balta, M. Kom
NIK. 190302290

Supriatna, M. Kom
NIK. 190302239



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 25 November 2022

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : M Zufar Taquuddin
NIM : 17.11.1304

Menyatakan bahwa Skripsi dengan judul berikut:

**IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK KLASIFIKASI
BUKU PERPUSTAKAAN PADA SISTEM INFORMASI PERPUSTAKAAN
KABUPATEN BANTUL.**

Dosen Pembimbing: Anna Baita, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 25 November 2022

Yang Menyatakan,

*Meterai Asli
Rp 10.000,-*

M Zufar Taquuddin

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : M Zufar Taquiuddin
NIM : 17.11.1304

Menyatakan bahwa Skripsi dengan judul berikut:

**IMPLEMENTASI ALGORITMA NAIVE BAYES UNTUK KLASIFIKASI
BUKU PERPUSTAKAAN PADA SISTEM INFORMASI PERPUSTAKAAN
KABUPATEN BANTUL.**

Dosen Pembimbing: Anna Baita, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 25 November 2022

Yang Menyatakan,



M Zufar Taquiuddin

HALAMAN PERSEMBAHAN

Alhamdulillah penulis panjatkan puji-syukur kepada Allah SWT atas segala rahmat, taufiq, serta hidayah-Nya, sehingga diberi kesempatan untuk dapat menyelesaikan skripsi ini dengan sebaik-baiknya dengan segala kekurangan penulis. Segala syukur penulis ucapkan kepada-Mu karena telah menghadirkan mereka yang memberikan semangat dan doa disaat menjalani proses pembuatan skripsi ini. Dengan segala kerendahan hati saya persembahkan skripsi ini kepada:

1. Kedua Orang Tua, Bapak Priyo Dwi Abang Suranto dan Ibu Adilah yang selalu mendoakan, memberi semangat serta motivasi supaya dapat menyelesaikan skripsi ini dengan lancar serta bermanfaat bagi semua.
2. Ibu Anna Baita ,M.Kom. selaku dosen pembimbing dalam skripsi ini yang tidak lelah untuk tetap membimbing dan mengingatkan penulis dari awal hingga akhir proses pembuatan skripsi.
3. Dosen-dosen Universitas Amikom Yogyakarta yang telah memberikan banyak ilmu baik ilmu akademik maupun ilmu non-akademik selama kuliah.
4. Serta orang-orang yang selalu membantu peneliti dalam mengerjakan skripsi yang tidak bisa disebut namanya satu-persatu.

Saya ucapkan terima kasih yang sebesar-besarnya untuk kalian semua. Sukses untuk kalian semua, semoga Allah SWT memberikan rahmat dan hidayah-Nya kepada kita semua. Dan semoga skripsi ini dapat bermanfaat dan berguna untuk kemajuan ilmu pengetahuan kedepannya.

KATA PENGANTAR

Alhamdulillah penulis panjatkan puji syukur kepada Allah SWT atas segala rahmat, taufiq, serta hidayah-Nya kepada penulis sehingga dapat menyelesaikan skripsi yang berjudul "Implementasi Algoritma Naïve Bayes untuk Klasifikasi Buku Perpustakaan pada Sistem Informasi Kabupaten Bantul".

Selama proses pengerjaan skripsi ini penulis menyadari bahwa dalam proses penulisan skripsi ini banyak mengalami kendala, namun berkat bantuan, bimbingan, kerjasama dari berbagai pihak dan berkah dari Allah SWT sehingga kendala-kendala tersebut bisa diatasi. Selanjutnya ucapan terima kasih penulis sampaikan kepada:

1. Kedua orang tua dan keluarga yang selalu memberikan doa, dukungan dan semangat
2. Ibu Anna Baita, M.Kom. selaku dosen pembimbing yang telah memberikan banyak masukan yang membantu membimbing dalam menyelesaikan skripsi ini
3. Terima kasih untuk seluruh Dosen Program Studi Informatika Amikom Yogyakarta yang telah memberikan ilmunya.
4. Perpustakaan Kabupaten Bantul yang telah bersedia untuk menjadi tempat penulis melakukan penelitian.
5. Terima kasih untuk teman satu lingkaran yang sudah sangat membantu dalam menyelesaikan skripsi ini.
6. Serta semua pihak yang tidak bisa penulis sebutkan satu-persatu yang telah membantu dalam penyusunan skripsi ini.

Penulis Menyadari bahwa masih banyak terdapat kekurangan-kekurangan dalam mengerjakan skripsi ini, sehingga penulis mengharapkan adanya saran dan kritik yang membangun demi kesempurnaan skripsi ini

Yogyakarta, 25 November 2022

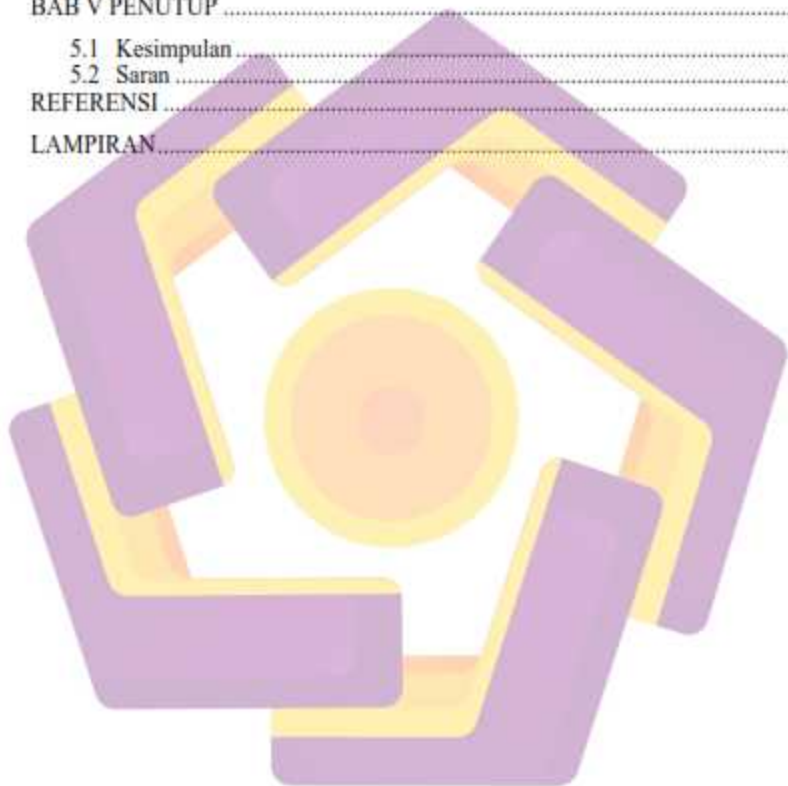
M Zufar Taqiuddin

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiv
INTISARI	xv
ABSTRACT	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	3
1.6 Metode Penelitian	4
1.6.1 Metode Pengumpulan Data	4
1.6.2 Metode Analisis	4
1.6.3 Metode Perancangan	5
1.7 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1 Studi Literatur	7
2.2 Keaslian Penelitian	8
2.3 Dasar Teori	13
2.3.1 Sentimen Analisis	13
2.3.2 Text Mining	13
2.3.3 Dewey Decimal Classification.....	13
2.3.4 Pengumpulan Data	14

2.3.5	Preprocessing	14
2.3.6	Indexing	15
2.3.7	Naïve Bayes	16
2.3.8	Multinomial Naive Bayes Classifier	17
2.3.9	Synthetic Minority Over-sampling Technique (SMOTE)	18
2.3.10	Django REST Framework	20
2.4	Metode Pengujian	20
2.4.1	Confusion Matrix	20
BAB III METODE PENELITIAN		22
3.1	Objek Penelitian	22
3.2	Alat dan Bahan	22
3.2.1	Hardware (Perangkat Keras)	22
3.2.1	Software (Perangkat Lunak)	22
3.3	Alur Penelitian	23
3.3.1	Pengumpulan Data	26
3.3.2	Preprocessing	26
3.3.3	Splitting Data	31
3.3.4	Indexing	31
3.3.5	Oversampling Data	42
3.3.6	Microservice Django	45
BAB IV HASIL DAN PEMBAHASAN		46
4.1	Dataset	46
4.2	Preprocessing	46
4.3	Indexing	52
4.4	Splitting Data Pengujian	53
4.5	Klasifikasi Naive Bayes	55
4.5.1	Confusion Matrix	56
4.5.2	Data Balancing	63
4.5.3	Analisa Hasil	70
4.6	Pemasangan Fitur	74

4.6.1	Joblib.....	74
4.6.2	Django Microservice.....	75
4.6.3	Ajax dan JQuery	81
4.6.4	Hasil Tampilan Antarmuka.....	84
BAB V PENUTUP		85
5.1	Kesimpulan.....	85
5.2	Saran	85
REFERENSI		86
LAMPIRAN.....		89



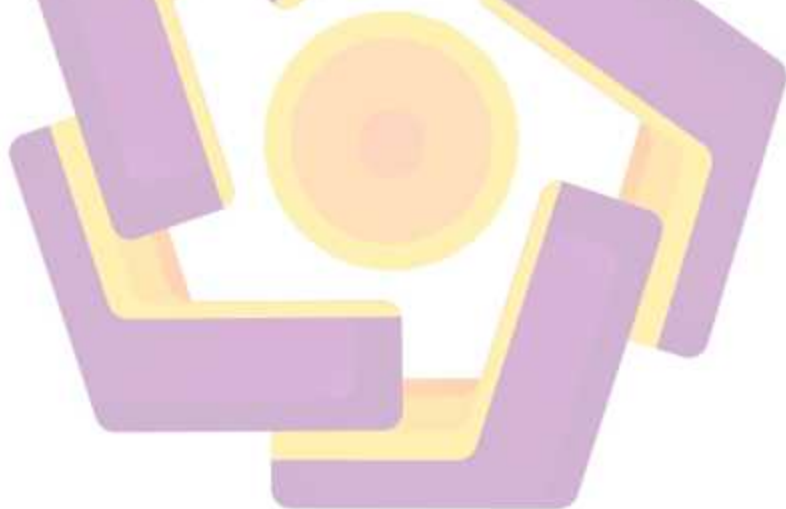
DAFTAR TABEL

Tabel 2. 1 Perbandingan dengan penelitian sebelumnya	9
Tabel 2. 2 Sepuluh kelas utama pada DDC	14
Tabel 2. 3 Confusion Matrix	20
Tabel 3. 1 Jumlah data judul pustaka perklasifikasi DDC	26
Tabel 3. 2 Tabel Data Hasil Preprocessing	30
Tabel 3. 3 Tabel spliting data yang akan di ujikan	31
Tabel 3. 4 Tabel Proses TF-IDF	32
Tabel 3. 5 Tabel Data Dari Proses Preprocessing	37
Tabel 3. 6 Tabe Hasil Proses TF-IDF	39
Tabel 3. 7 Tabel Proses Preprocessing	40
Tabel 3. 8 Tabel Proses Preprocessing data uji Naive Bayes	41
Tabel 3. 9 Tabel Judul Buku Pada Klasifikasi Bahasa	42
Tabel 3. 10 Tabel Judul Buku Pada Klasifikasi Bahasa	43
Tabel 4. 1 Tabel Data Hasil Preprocessing	51
Tabel 4. 2 Tabel spliting data yang akan di ujikan	53
Tabel 4. 3 Tabel hasil akurasi pada variasi spliting data	63
Tabel 4. 4 Tabel hasil akurasi pada variasi spliting data over sampling	70
Tabel 4. 5 Tabel analisa hasil	73

DAFTAR GAMBAR

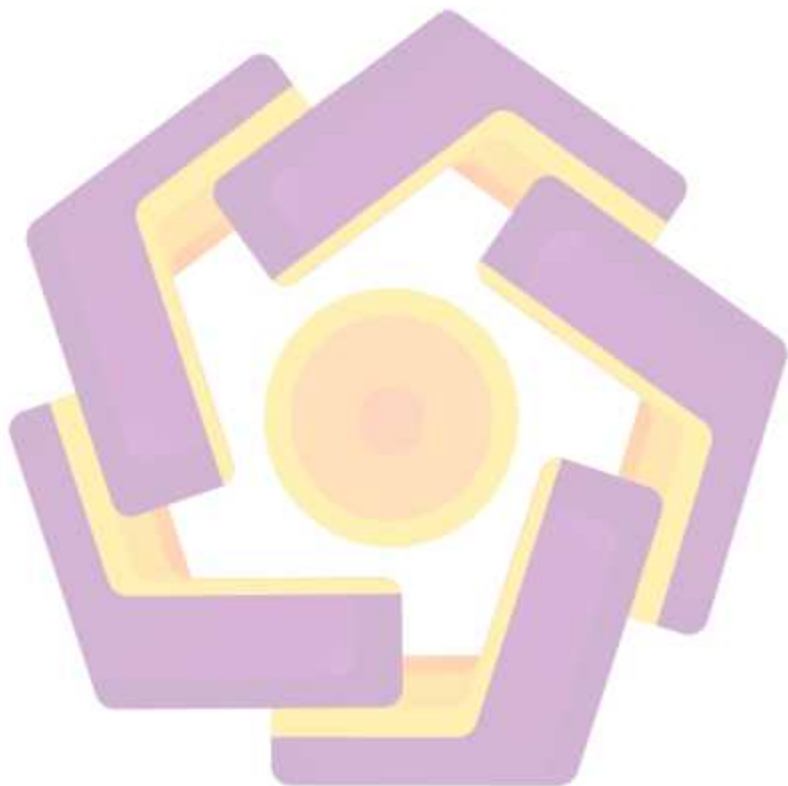
Gambar 2. 1 Ilustrasi Distribusi Data Metode SMOTE.....	19
Gambar 3. 1 Alur Penelitian	26
Gambar 3. 2 Alur Diagram Preprocessing	27
Gambar 3. 3 Alur Diagram <i>Stopword Removal</i>	28
Gambar 3. 4 Alur Diagram Stemming	29
Gambar 3. 5 Ilustrasi Spliting Data	31
Gambar 3. 6 Proses Klasifikasi Teks	33
Gambar 3. 7 Proses training Pembentukan Model Pengklasifikasi	35
Gambar 3. 8 Proses testing Pembentukan Model Pengklasifikasi	37
Gambar 3. 9 Alur Diagram Model Naive Bayes pada Django.	45
Gambar 4. 1 Visualisasi jumlah data	46
Gambar 4. 2 Kode program proses <i>case folding</i>	47
Gambar 4. 3 Gambaran hasil proses Tokenizing	47
Gambar 4. 4 Kode program proses Tokenizing	48
Gambar 4. 5 Kode program proses <i>s Stopward Removal</i>	49
Gambar 4. 6 Hasil proses Stopward Removal	50
Gambar 4. 7 Kode program proses Stemming	50
Gambar 4. 8 Gambaran hasil proses Stemming	51
Gambar 4. 9 Kode program TF-IDF	52
Gambar 4. 10 Memeriksa kode index TF-IDF	52
Gambar 4. 11 Memeriksa index term dengan array TF-IDF	53
Gambar 4. 12 Kode program <i>spliting data</i>	54
Gambar 4. 13 Kode program model Multinomial Naive Bayes	55
Gambar 4. 14 Kode program menampilkan confusion matrix	56
Gambar 4. 15 Percobaan 75:25 Confusion Matrix	57
Gambar 4. 16 Hasil pengukuran percobaan 75:25	58
Gambar 4. 17 Percobaan 80:20 Confusion Matrix	59
Gambar 4. 18 Hasil pengukuran percobaan 80:20	59
Gambar 4. 19 Percobaan 85:15 Confusion Matrix	60
Gambar 4. 20 Hasil pengukuran percobaan 85:15	61
Gambar 4. 21 Percobaan 90:10 Confusion Matrix	62
Gambar 4. 22 Hasil pengukuran percobaan 90:10	62
Gambar 4. 23 Kode Program SMOTE	63
Gambar 4. 24 Percobaan 75:25 Confusion Matrix	64
Gambar 4. 25 Hasil pengukuran percobaan 75:25	65
Gambar 4. 26 Percobaan 80:20 Confusion Matrix	66
Gambar 4. 27 Hasil pengukuran percobaan 80:20	66
Gambar 4. 28 Percobaan 85:15 Confusion Matrix	67
Gambar 4. 29 Hasil pengukuran percobaan 85:15	68

Gambar 4. 30 Percobaan 90:10 Confusion Matrix	69
Gambar 4. 31 Hasil pengukuran percobaan 90:10	69
Gambar 4. 32 Confusion Matrix sebelum di oversampling 80:20	70
Gambar 4. 33 Confusion Matrix sesudah di oversampling 80:20	71
Gambar 4. 34 data training sebelum dan sesudah oversampling	71
Gambar 4. 35 precision, recall, dan f12-score sebelum oversampling 80:20	71
Gambar 4. 36 precision, recall, dan f12-score sesudah oversampling 80:20	72
Gambar 4. 37 Kode Program Implementasi Joblib	75
Gambar 4. 38 Struktur dokumen Program Django Microservice	76
Gambar 4. 39 File views.py pada folder api	77
Gambar 4. 40 Fungsi NBTtitle pada file views.py pada folder api	78
Gambar 4. 41 File urls.py pada folder api	79
Gambar 4. 42 Hasil Response URL Judul	80
Gambar 4. 43 Tag HTML Judul, Subyek, dan Klasifikasi	81
Gambar 4. 44 Kode Javascript	82
Gambar 4. 45 Kode AJAX	83
Gambar 4. 46 Tampilan Halaman Pengimputan Koleksi Buku	84



DAFTAR LAMPIRAN

Lampiran 1. Profil obyek Penelitian	10
Lampiran 2. Dokumentasi Penelitian	11



INTISARI

Teknologi informasi di era revolusi industri 4.0 menuntut pustakawan untuk mampu berkolaborasi dan berkomunikasi secara baik yang telah terotomasi kedalam suatu sistem. Kinerja seorang pustakawan perlu dipertimbangkan dari kompetensi atau kemampuan yang dimiliki dan menyesuaikan diri dengan perkembangan teknologi informasi saat ini. Pekerjaan yang membutuhkan fokus dan ketelitian pada setiap layanan terkadang menemui sebuah permasalahan yang dapat menghambat kinerja seorang pustakawan. Dengan teknologi informasi dan komunikasi baik aplikasi maupun platform telah merubah Pustakawan dapat melakukan kolaborasi dengan media teknologi informasi yang saat ini dikenal dengan istilah kecerdasan buatan atau artificial intelligence (AI). Penggunaan algoritma AI dalam beragam perangkat lunak komputer membantu manusia dalam melakukan pekerjaannya secara otomatis saahsatunya dalam pengklasifikasian data.

Metode yang digunakan untuk klasifikasi data judul buku adalah algoritma naïve bayes dengan tahap evaluasi performasi menggunakan metode Confusion Matrix. Menggunakan dataset dari Perpustakaan Daerah Kabupaten Bantul dengan hasil fitur pengklasifikasian buku otomatis pada proses input judul koleksi.

Fitur ini dapat digunakan oleh pustakawan Perpustakaan Daerah Kabupaten Bantul untuk memudahkan proses pengolahan bahan pustaka dalam melakukan pengklasifikasian judul buku

Kata kunci: Naïve Bayes, Klasifikasi, Kecerdasan Buatan.

ABSTRACT

Information technology in the era of the industrial revolution 4.0 requires librarians to be able to collaborate and communicate well that has been automated into a system. The performance of a librarian needs to be considered from the competencies or abilities possessed and to adapt to current developments in information technology. Jobs that require focus and thoroughness in each service sometimes encounter a problem that can hinder a librarian's performance. With information and communication technology both applications and platforms, librarians can collaborate with information technology media which is currently known as artificial intelligence (AI). The use of AI algorithms in various computer software helps humans to do their work automatically, one of which is in classifying data.

The method used to classify book title data is the naive Bayes algorithm with the performance evaluation stage using the Confusion Matrix method. Using a dataset from the Bantul Regency Regional Library with the results of the automatic book classification feature in the collection title input process.

This feature can be used by the librarian of the Bantul Regency Regional Library to facilitate the process of processing library materials in classifying book titles.

Keyword: Naive Bayes, Classification, Artificial Intelligence