

Rancang Bangun Sistem Klasifikasi Kategori Buku dengan Algoritma *Naïve Bayes Classifier* (Studi Kasus : Perpustakaan Universitas Dinamika)

Casmina ¹⁾ Vivine Nurcahyawati ²⁾ Julianto Lemantara

Program Studi/Jurusan Sistem Informasi

Universitas Dinamika

Jl. Raya Kedung Baruk 98 Surabaya, 60298

Email : 1)17410100191@dinamika.ac.id, 2)vivine@dinamika.ac.id, 3)julianto@dinamika.ac.id

Abstract: *Dinamika University library has an online catalog to search for books. to determine the category / class referring to the ddc23 book takes approximately 3 to 10 minutes. the problem that exists is that the library manager has difficulty determining the appropriate category because the determination of the category is not fast enough. the purpose of this research is to produce a book category classification system with the naïve bayes classifier algorithm (case study: undika library) to determine classification quickly and improve the suitability of book categories. naïve bayes classifier is a method used for classification in which the process is divided into two stages, namely the learning stage and the testing phase. classification uses class 000 of ddc23, where the data is divided into 311 training data and 78 testing data and the results of the accuracy test are 79% which shows the naïve bayes classifier algorithm can be used as an algorithm for classifying book categories because of its high accuracy. the level of testing accuracy is influenced by the large amount of training data and the category dictionary used. Classification of categories is done in a faster time with the system only takes 2 seconds 70 milliseconds.*

Keywords: *Text classification, Naïve Bayes Classifier, Text Preprocessing, Book category.*

Perpustakaan adalah tempat dimana banyaknya koleksi pustaka disimpan dan disediakan untuk pemustaka dengan pelayanan dan pengelolaan pustaka yang baik. Seperti dalam penelitian ini yaitu Perpustakaan Universitas Dinamika (Undika) mempunyai tempat yang nyaman dan fasilitas lengkap untuk kebutuhan pemustaka. Pada segi kemampuan teknologi Perpustakaan Undika mempunyai *Online Catalog* untuk melakukan pencarian buku. Untuk mendaftarkan koleksi ke aplikasi pengadaan dilakukan oleh kepala bagian perpustakaan. Kemudian penentuan kategori/kelas oleh petugas pengelola pustaka yang mengacu pada buku DDC23 mulai dari pencarian pada kelas utama ke kelas kedua dan kelas ketiga sebagai kelas spesifik atau kelas kecil dari penentuan kategori tersebut. Dilanjutkan ke pengolahan untuk menentukan data dengan pengindukan atau penentuan klasifikasi yang membutuhkan waktu kurang lebih 3 sampai 10 menit dilanjutkan input deskripsi, setelah klik simpan maka data buku akan masuk ke sistem *Online Catalog* dan dapat diakses secara umum.

Permasalahan yang ada yaitu pengelola pustaka kesulitan menentukan kategori yang sesuai karena penentuan kategori kurang cepat, dilakukan berdasarkan klasifikasi (nomor) yang

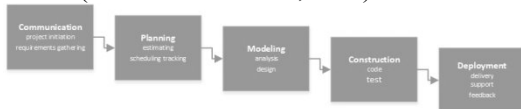
ada pada DDC dari mulai mencari pada klasifikasi besar hingga pencarian spesifik ke klasifikasi kecil sampai menemukan kategori yang sesuai. Dampak yang ditimbulkan yaitu penentuan kategori memakan waktu dan memungkinkan terjadinya kesalahan pengklasifikasian kategori buku yang akan mempengaruhi pencarian katalog pada saat digunakan oleh pemustaka.

Solusi yang digunakan untuk membantu mempermudah pengelola dalam melakukan klasifikasi kategori yang sesuai yaitu dengan merancang sistem klasifikasi kategori buku dengan algoritma *Naïve Bayes Classifier* (Studi Kasus : Perpustakaan Undika) untuk penentuan klasifikasi secara cepat dan meningkatkan kesesuaian kategori buku. Kriteria klasifikasi berasal dari kelas 000 berdasarkan DDC23 sejumlah 16 meliputi *Computer Program Language, Computer Programs, Operating Systems, Computers, Computer Network, Computer Security, Computer Science, Internet, Computer Graphics, Library, Operations research, Microprocessors, Artificial Intelligence, System Analysis, Data Communication, Microcomputers*. Jumlah data yang digunakan sebanyak 389 buku yang terdiri dari 80% data pelatihan dan 20% data pengujian.

Tujuan dari penelitian ini yaitu Menghasilkan sistem klasifikasi kategori buku dengan algoritma *Naïve Bayes Classifier* (Studi Kasus : Perpustakaan Undika) untuk penentuan klasifikasi secara cepat dan meningkatkan kesesuaian kategori buku. Landasan teori penelitian menggunakan Penelitian Terdahulu, *Data mining*, klasifikasi, Algoritma *Naïve Bayes Classifier*, dan *Text Preprocessing*. Dua penelitian terdahulu dari (Indriyono, Utami, & Sunyoto, 2015) menggunakan 9 kategori dengan jumlah 900 data sample dan (Rizqiyani, Mulwida, & Putri, 2017) menggunakan 6 kategori dengan jumlah 249 data sample. Sementara penelitian saat ini akan merancang sistem klasifikasi kategori buku dengan algoritma *Naïve Bayes Classifier* (Studi Kasus : Perpustakaan Undika) untuk penentuan klasifikasi secara cepat dan meningkatkan kesesuaian kategori buku. kriteria klasifikasi berasal dari kelas 000 berdasarkan DDC23. Klasifikasinya ada 16 kategori yaitu *Computer Program Language, Computer Programs, Operating Systems, Computers, Computer Network, Computer Security, Computer Science, Internet, Computer Graphics, Library, Operations research, Microprocessors, Artificial Intelligence, System Analysis, Data Communication, Microcomputers*. Dan data yang digunakan berasal dari koleksi Perpustakaan Undika. Jumlah data yang digunakan sebanyak 389 buku yang terdiri dari 80% data pelatihan dan 20% data pengujian.

METODE

Model air terjun, kadang-kadang disebut siklus hidup klasik, menyarankan pendekatan sistematis, sekuensial 2 untuk pengembangan perangkat lunak yang dimulai dengan spesifikasi kebutuhan pelanggan dan berkembang melalui perencanaan, pemodelan, konstruksi, dan penerapan, yang berpuncak pada dukungan berkelanjutan dari perangkat lunak yang sudah selesai (Pressman & Maxim, 2015).



Gambar 1 SDLC model *Waterfall* Menurut (Pressman & Maxim, 2015)

1. Communication

Pada tahap awal ini peneliti melakukan observasi pada Perpustakaan Undika dan melakukan wawancara langsung dengan pihak Perpustakaan Undika yaitu Ibu Deasy Kumalawati sebagai Kepala Bagian Perpustakaan dan Bapak Totok sebagai petugas pengelolaan pustaka.

permasalahan yang ada yaitu pengelola pustaka kesulitan menentukan kategori yang sesuai karena penentuan kategori kurang cepat. Klasifikasi dilakukan berdasarkan kelas yang ada pada DDC dari mulai mencari pada klasifikasi besar hingga pencarian ke klasifikasi kecil sampai menemukan kategori yang sesuai.

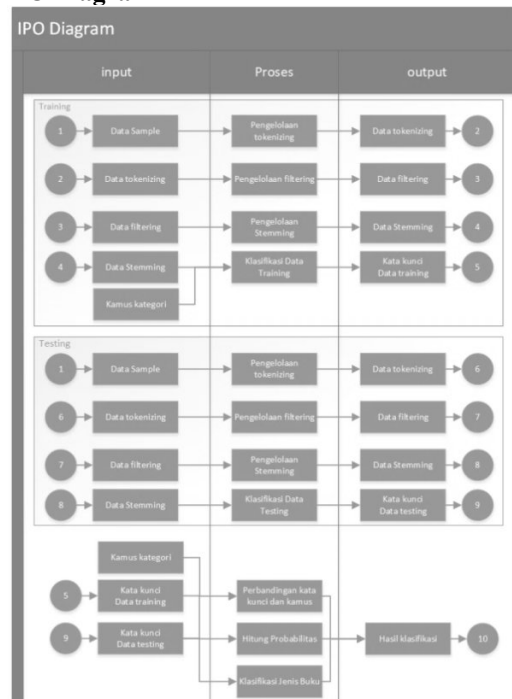
2. Planning

Pada tahap ini akan dilakukan penjadwalan untuk membuat rincian waktu pertahap yang akan dikerjakan. Untuk gambaran yang lebih jelas akan ditunjukkan pada Jadwal Kerja.

3. Modeling

Tahap ini digunakan untuk mengubah kebutuhan - kebutuhan diatas menjadi representasi sebelum coding dimulai. Mengimplementasikan kebutuhan yang telah disebutkan pada tahap sebelumnya.

IPO Diagram



Gambar 2 IPO Diagram

Berikut Penjelasan dari Gambar IPO diagram diatas :

Input

1. *Data Sample*
Merupakan data yang digunakan untuk implementasi yang terbagi menjadi data *training* dan data *testing*.
2. *Data tokenizing*
Berisi data hasil dari proses *tokenizing* yang mengubah bentuk kalimat ke daftar kata-kata.
3. *Data filtering*

Berisi daftar kata-kata penting yang sudah dicocokkan untuk membuang kata-kata yang tidak diperlukan atau kurang penting menggunakan *Stopword/Stoplist*.

4. Data *stemming*

Berisi daftar kata-kata dasar yang telah diubah dari *filtering* yang akan digunakan untuk kata kunci.

5. Data *Testing*

Sebagai data yang digunakan untuk proses perhitungan klasifikasi buku. Berisi judul atau deskripsi yang akan dimasukan.

6. Kamus Kategori

Berisi kumpulan kata-kata dari kategori tertentu.

Proses

1. Pengelolaan *Tokenizing*

Proses ini dilakukan untuk mengubah bentuk kalimat ke satuan kata-kata.

2. Pengelolaan *filtering*

Bagian ini adalah untuk pengelolaan dari *tokenizing* menjadi kata-kata deskriptif seperti “yang”, “dan”, dan “di” dengan *Stopword* yang digunakan untuk proses *stemming*.

3. Pengelolaan *Stemming*

Proses ini akan mengubah data dari *filtering* ke bentuk kata dasar.

4. Klasifikasi Data *Training*

Pada proses ini dilakukan klasifikasi menggunakan kamus kategori.

5. Klasifikasi Data *Testing*

Pada proses ini dilakukan klasifikasi untuk menghilangkan teks *noise* untuk digunakan pada klasifikasi perbandingan dengan data *training* dan kamus.

6. Perbandingan kata kunci dan kamus

Merupakan proses pemilihan keakuratan kata disesuaikan dengan data buku pengujian dan kamus.

7. Hitung probabilitas

Proses perhitungan probabilitas untuk menentukan klasifikasi.

8. Hitung Klasifikasi

Proses ini yaitu menghitung klasifikasi sesuai rumus NBC untuk menentukan klasifikasi kategori yang sesuai.

Output

1. Data *Tokenizing*

Berisi data hasil dari proses *tokenizing* yang mengubah bentuk kalimat ke daftar kata-kata.

2. Data *Filtering*

Berisi daftar kata-kata penting yang sudah dicocokkan untuk membuang kata-kata yang tidak diperlukan atau kurang penting menggunakan *Stopword/Stoplist*.

3. Data *Stemming*

Berisi daftar kata-kata dasar yang telah diubah dari *filtering* yang akan digunakan untuk kata kunci.

4. Kata kunci data *Training*

Berisi data kata-kata yang akan dilakukan perbandingan pada tahap *Testing*.

5. Kata kunci data *Testing*

Berisi data kata-kata yang akan dilakukan perbandingan dengan hasil kata kunci pada tahap *Training*.

6. Hasil klasifikasi

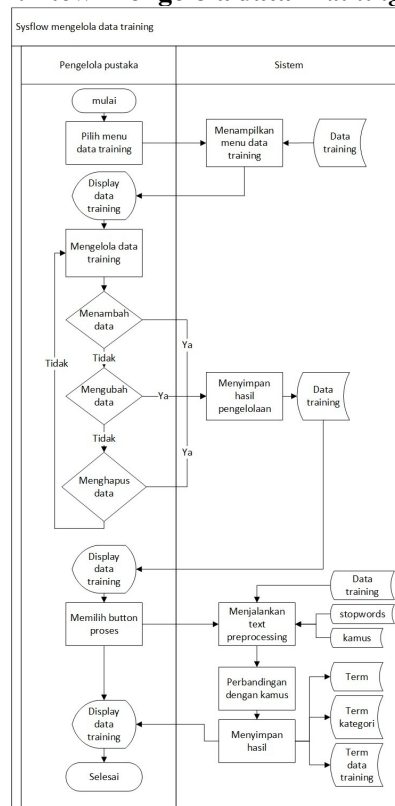
Hasil ini menunjukkan kategori yang sesuai dengan buku yang telah dilakukan *testing* klasifikasi dengan NBC.

Desain Sistem

Berikut ini merupakan rancangan sistem klasifikasi kategori.

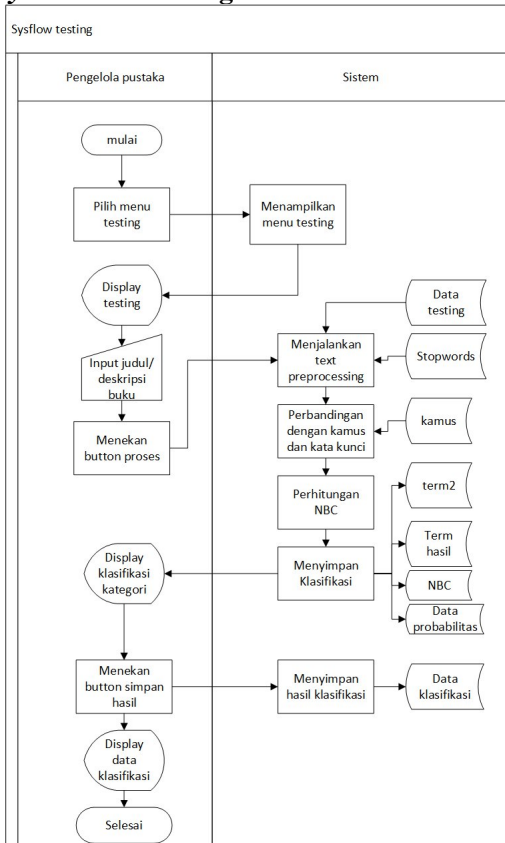
System Flow

System Flow Mengelola data *Training*



Gambar 3 System Flow Mengelola data *Training*
 Terlihat pada gambar diatas menggambarkan tentang alur sistem untuk mengelola data *training* dan menjalankan rangkaian proses *text preprocessing* seperti *tokenizing*, *filtering*, dan *stemming* kemudian dilakukan perbandingan dengan kamus kategori yang selanjutnya digunakan untuk proses pengujian.

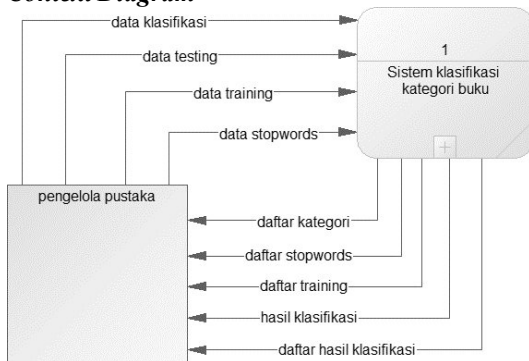
Sytem Flow Testing



Gambar 4 System Flow Testing

Pada Gambar diatas menunjukkan alur sistem melakukan pengujian data menggunakan rangkaian proses *text preprocessing* yaitu *tokenizing*, *filtering*, dan *stemming* dan menjalankan penghitungan NBC.

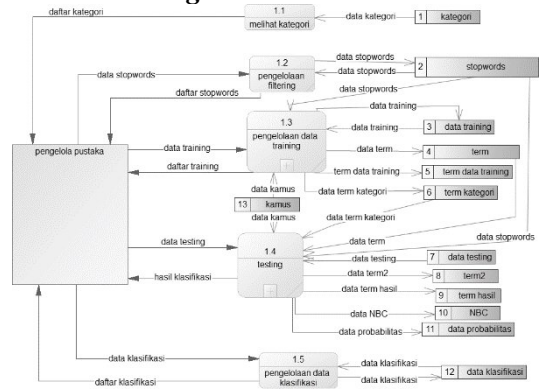
Data Flow Diagram Context Diagram



Gambar 5 Context Diagram sistem

Gambar diatas merupakan *Context Diagram* untuk menunjukkan hubungan entitas dan proses sistem. Pada proses sistem terdapat *decompose process* yang mana prosesnya akan dilanjutkan pada *Data Flow Diagram level 0*.

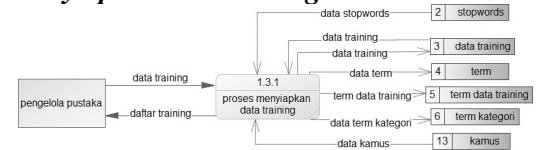
Data Flow Diagram level 0



Gambar 6 Data Flow Diagram level 0

Gambar diatas menunjukkan hubungan entitas (simbol kotak berwarna biru tua) dengan setiap proses yang ada (simbol kotak berwarna biru muda dengan nomor 1.1 sampai 1.5) pada sistem dengan database (simbol berwarna kuning).

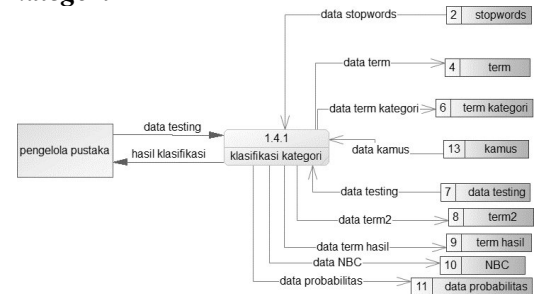
Data Flow Diagram level 1 proses menyiapkan data training



Gambar 7 Data Flow Diagram level 1 proses menyiapkan data training

Pada gambar diatas menunjukkan proses *decompose* dari proses pengelolaan data *training*.

Data Flow Diagram level 1 klasifikasi kategori



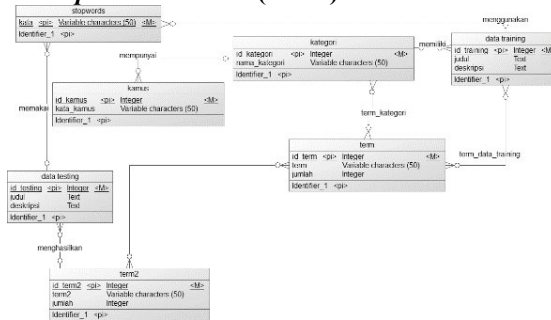
Gambar 8 Data Flow Diagram level 1 klasifikasi kategori

Pada gambar diatas merupakan *decompose* proses dari proses testing pada *Data Flow Diagram level 0*.

Merancang Basis data

Dibawah ini adalah rancangan skema *database* yang akan digunakan sistem. Ditunjukan dalam bentuk *Conceptual data model*, *physical data model*, dan struktur tabel.

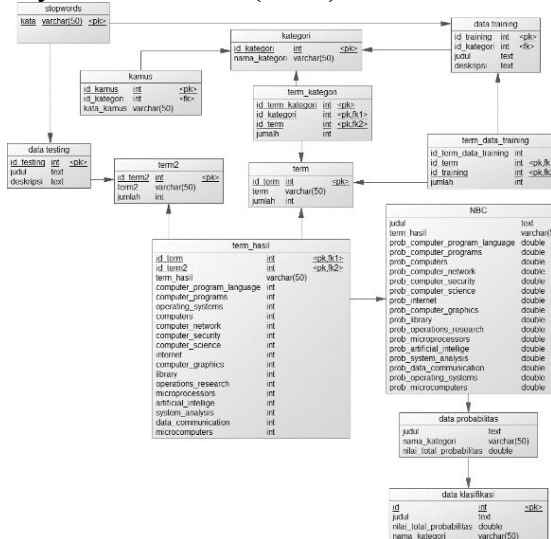
Conceptual data model(CDM)



Gambar 9 Conceptual data model Sistem

Pada Gambar diatas menunjukkan entitas yang dihubungkan dengan entitas lain membentuk relasi. Pada masing masing entitas juga ditunjukkan atribut, tipe data, dan masing - masing panjangnya.

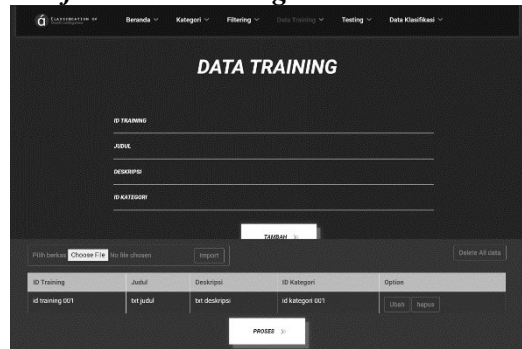
Physical data model (PDM)



Gambar 10 Physical data model Sistem

Dari gambar diatas menunjukkan entitas – entitas generate dari CDM. Terdapat entitas baru yang dihasilkan karena relasi yang dibuat berdasarkan *cardinalitas many to many* dan terdapat tabel baru yang dibuat setelah muncul tabel hasil relasi untuk menyimpan kebutuhan data tambahan yaitu NBC, data probabilitas, dan data klasifikasi.

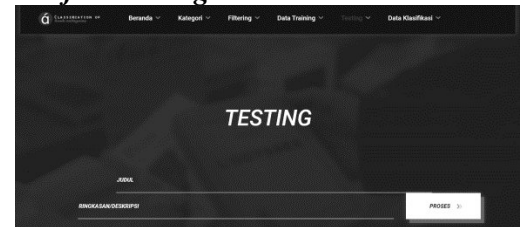
Desain Interface Interface data Training



Gambar 11 Interface data training

Gambar diatas merupakan tampilan untuk memasukan data training. Dengan desain konsisten memakai *background* kontras kombinasi warna ungu dan biru. Daftar data training ditampilkan rapih berbentuk tabel dibuat transparan dengan button berwarna biru.

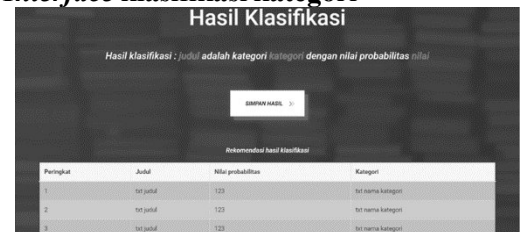
Interface Testing



Gambar 12 Interface Testing

Pada Gambar diatas menunjukkan form input yang di desain menarik mata pengguna, font italic dan *background* transparan blur dengan gambar seseorang memegang buku diantara kumpulan buku.

Interface klasifikasi kategori



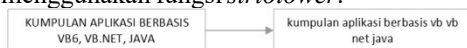
Gambar 13 Interface klasifikasi kategori

Tampilan yang ditunjukkan pada Gambar diatas mempunyai desain yang sama dengan halaman sebelumnya agar terlihat serasi dan kombinasi yang bagus dari warna tabel putih dan warna background gradient biru dan ungu.

4. Construction

1. *Case Folding* yaitu mengubah semua huruf dalam teks data *Training* menjadi huruf kecil. Hanya huruf 'a' sampai dengan 'z' yang diterima. Karakter selain huruf dihilangkan

dan dianggap delimiter. Kemudian dilanjutkan ke tahap selanjutnya yaitu *Tokenizing*. Untuk membuat judul dan deskripsi buku hanya huruf a sampai z menggunakan fungsi *strtolower*.



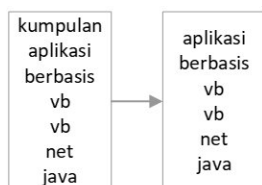
Gambar 14 proses Case folding

- Tokenizing* adalah sebuah proses untuk memilah isi teks sehingga menjadi satuan kata-kata. Proses ini mengambil data dari hasil *Case Folding* untuk membuat kata per kata menggunakan fungsi *str replace*.



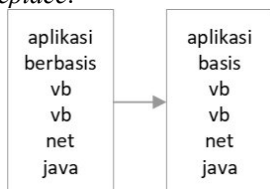
Gambar 15 proses *Tokenizing*

- Filtering* merupakan tahap mengambil kata-kata penting dari hasil *token*. menggunakan algoritma *Stoplist* (membuang kata yang kurang penting) atau *wordlist* (menyimpan kata penting). Menggunakan perulangan *for* dan pengkondisian *if else* untuk memilih hanya kata – kata yang selain dari data *filtering* yang sudah disiapkan menggunakan tabel.



Gambar 16 proses *filtering*

- Stemming* merupakan suatu proses untuk mereduksi kata ke bentuk dasarnya untuk menghilangkan imbuhan – imbuhan yang ada pada setiap kata menggunakan fungsi *preg_replace*.



Gambar 17 proses *Stemming*

5. Deployment

Skenario pengujian

Pada Sistem klasifikasi kategori ini terdapat dua pengujian yaitu *Black Box* untuk menguji fungsional sistem dan pengujian akurasi.

Pengujian *Black Box*

Berikut ini merupakan daftar fungsi – fungsi yang akan diuji lengkap dengan *Test Case* yang digunakan. Untuk fungsi pendukung dapat dilihat pada halaman lampiran.

Tabel 1 *Test Case*

No	Fungsi	Test Case
1	Proses klasifikasi data Training	TC0701
		TC0702
		TC0703
2	Testing	TC0801
		TC0802
		TC0803

Proses klasifikasi data Training

Tabel 2 hasil pengujian proses klasifikasi data Training

Test Case	Expected Result	Result
TC0701	Berhasil, dan muncul <i>pop up</i> klasifikasi data training selesai.	Berhasil
TC0702	Gagal, data Training kosong.	
TC0703	Berhasil, dan kembali ke halaman Training	Berhasil

Testing

Tabel 3 hasil pengujian Testing

Test Case	Expected Result	Result
TC0801	Berhasil, dan muncul <i>pop up</i> klasifikasi data training selesai.	Berhasil
TC0802	Gagal, inputan testing tidak terisi.	Berhasil
TC0803	Berhasil, dan menampilkan halaman klasifikasi menunjukkan hasil klasifikasi kategori dari data Testing.	Berhasil

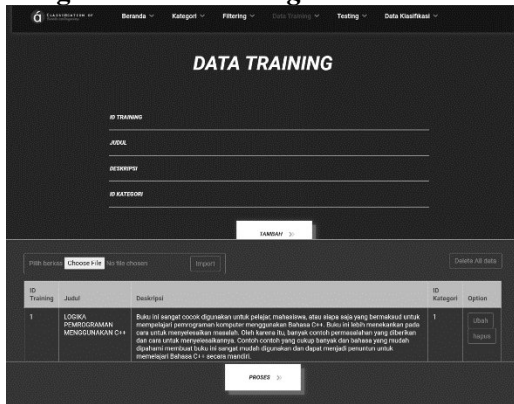
Pengujian Akurasi

Pengujian Akurasi klasifikasi kategori dilakukan dengan membandingkan klasifikasi kategori dari Perpustakaan Undika dan klasifikasi kategori yang dilakukan menggunakan sistem dengan algoritma *Naive Bayes Classifier*. Pengujian menggunakan 311 data Training dan 78 data Testing.

HASIL DAN PEMBAHASAN
Implementasi Sistem

Tahap implementasi sistem akan memberikan penjelasan dari rancangan sistem yang sudah dibuat untuk menunjukkan bagaimana cara kerja Sistem Klasifikasi Kategori pada Perpustakaan Undika. Fungsi pendukung lainnya dapat dilihat pada halaman lampiran.

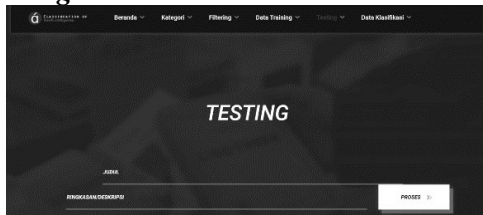
Mengelola data Training



Gambar 18 Halaman Data Training

Halaman pada Gambar diatas merupakan untuk mengelola data *Training* yaitu tambah, ubah, hapus, dan menampilkan data *Training*. *Button* proses yang berada di bagian bawah merupakan *Button* untuk melakukan klasifikasi data *Training*.

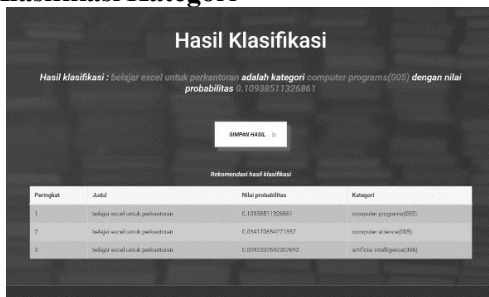
Testing



Gambar 19 halaman Testing

Gambar diatas merupakan halaman input data testing berisi judul dan deskripsi buku yang akan dilakukan *Testing* klasifikasi kategori, data akan diproses setelah pengguna yaitu pengelola pustaka menekan *Button* proses.

Klasifikasi Kategori



Gambar 20 halaman klasifikasi Kategori

Pada Gambar diatas menunjukkan halaman hasil klasifikasi kategori yang diproses oleh sistem sesuai tahap pertahap yang sudah dirancang dan klasifikasi menggunakan algoritma *Naïve Bayes Classifier*.

Pengujian Sistem

Dibawah ini merupakan pengujian sistem yang dilakukan untuk mengetahui fungsi – fungsi yang ada telah sesuai dengan kebutuhan dan sesuai fungsinya.

Proses klasifikasi data Training

Tabel 4 hasil pengujian proses klasifikasi data *Training*

Test Case	Expected Result	Result
TC0701	Berhasil, dan muncul <i>pop up</i> klasifikasi data training selesai.	Berhasil
TC0702	Gagal, data Training kosong.	
TC0703	Berhasil, dan kembali ke halaman <i>Training</i>	Berhasil

Testing

Tabel 5 hasil pengujian *Testing*

Test Case	Expected Result	Result
TC0801	Berhasil, dan muncul <i>pop up</i> klasifikasi data training selesai.	Berhasil
TC0802	Gagal, inputan testing tidak terisi.	Berhasil
TC0803	Berhasil, dan menampilkan halaman klasifikasi hasil klasifikasi kategori dari data <i>Testing</i> .	Berhasil

Hasil Pengujian

Tabel 6 Hasil pengujian

No	Judul	Klasifikasi oleh Perpustakaan Undika	Kategori sistem
1	membuat aplikasi database dengan java 2: panduan aplikatif	Computer program language	computer program language
2	membuat laporan pdf berbasis web dengan php 5.0	Computer program language	computer program language
3	pemrograman c dan c++	Computer program language	computer program language
4	tuntunan pemrograman java ,jilid 3	Computer program language	data communication
5	pemrograman client server microsoft visual basic 6.0, jilid 2	Computer program language	computer program language
6	pengolahan database mysql 5 dengan java 2 disertai teknik pencetakan laporan	Computer program language	computer program language
7	pengembangan aplikasi sistem informasi akademik berbasis sms dengan java	Computer program language	computer program language
8	bank soal visual basic	Computer program language	Computer program language
9	petunjuk praktis membangun aplikasi php point of sale	Computer program language	computer program language
10	web database menggunakan mysql dan sqlite	Computer program language	computer program s

No	Judul	Klasifikasi oleh Perpustakaan Undika	Kategori sistem
11	visual basic: membuat animasi dan tampilan cantik pada interface form	Computer program language	computer program s
12	pemrograman visual dengan c++ builder	Computer program language	computer program language
13	aplikasi excel dalam mengolah data transaksi dan laporan penjualan	computer program s	computer program s
14	cara mudah menggunakan spss data entry	computer program s	computer program s
15	aplikasi excel dalam pivottable bisnis terapan	computer program s	computer program s
16	panduan lengkap spss 13.0 dalam mengolah data statistik	computer program s	computer program s
17	desain aplikasi penggajian dan pph pasal 21 menggunakan microsoft access	computer program s	computer program s
18	aplikasi excel untuk analisis keuangan	computer program s	computer program s
19	cara mudah menggunakan crystal reports xi: pembuat laporan dari berbagai format data lokal dan sever	computer program s	computer program language
20	strategi jitu memilih metode statistik penelitian dengan spss	computer program s	computer program s

No	Judul	Klasifikasi oleh Perpustakaan Undika	Kategori sistem
21	utility windows password recovery	computer program s	computer security
22	microsoft visual foxpro 9.0: seri panduan lengkap	computer program s	computer program s
23	microsoft excel 2003, edisi kedua: buku materi kuliah stikom	computer program s	computer program s
24	microsoft acces 2003	computer program s	computer program s
25	7 jam belajar autocad 2006 untuk orang awam	computer program s	artificial intelligence
26	mengolah data spasial dengan mapinfo professional	computer program s	computer program s
27	microsoft excel 2000 for windows 98/2000 : belajar sendiri, edisi baru	computer program s	computer program s
28	sistem operasi microsoft windows 2000 professional : belajar sendiri, edisi baru	computer program s	computer network
29	buku mini info linux redhat 9	Operating systems	operating systems
30	linux : singkat tepat jelas	Operating systems	computer network
31	sistem operasi, edisi ke2	Operating systems	operating systems
32	panduan belajar sistem operasi	Operating systems	operating systems
33	mandriva linux, limited edition 2005	Operating systems	operating systems
34	prosiding seminar nasional ilmu komputer dan aplikasinya (snika) 2006, volume 1, no.1, november 2006	Computers	computer s
35	pc troubleshooting plus, edisi revisi kedua	Computers	Computers
36	18 program jitu untuk meningkatkan kinerja komputer	Computers	computer science
37	kamus lengkap jaringan komputer	computer network	computer network
38	multihoming menggunakan bgp (border gateway protocol) : membangun multi-koneksi ke multi-isp	computer network	computer network
39	jaringan komputer : buku materi kuliah stikom, edisi pertama	computer network	computer network
40	pengelolaan sistem jaringan : modul praktikum	computer network	computer network
41	membangun warnet dan intranet dengan teknologi diskless thin-client linux terminal server project versi 3	computer network	operating systems

No	Judul	Klasifikasi oleh Perpustakaan Undika	Kategori sistem	No	Judul	Klasifikasi oleh Perpustakaan Undika	Kategori sistem
42	praktikum <i>microsoft windows server</i> 2003 mengkonfigurasi jaringan dan <i>web server</i>	<i>computer network</i>	<i>computer network</i>	54	uml distilled, edisi 3: panduan singkat bahasa pemodelan objek standar	<i>computer science</i>	<i>computer science</i>
43	computer networking first step	<i>computer network</i>	<i>computer network</i>	55	aneka tool pembersih hard disk	<i>computer science</i>	<i>computer security</i>
44	konsep jaringan komputer dan pengembangannya	<i>computer network</i>	<i>computer network</i>	56	mikroprosesor dan <i>interfacing</i>	<i>computer science</i>	<i>operating systems</i>
45	panduan membangun jaringan komputer untuk pemula	<i>computer network</i>	<i>computer network</i>	57	facebook	<i>internet</i>	<i>internet</i>
46	meningkatkan kemampuan jaringan komputer dengan pc <i>cloning system</i>	<i>computer network</i>	<i>computer network</i>	58	search engine <i>optimization</i> : cara cepat mendapatkan rating tinggi di search engine	<i>internet</i>	<i>internet</i>
47	kamus++ jaringan komputer	<i>computer network</i>	<i>computer network</i>	59	google cheat teknik jitu	<i>internet</i>	<i>internet</i>
48	jaringan komputer	<i>computer network</i>	<i>computer network</i>	60	memfilter internet	<i>internet</i>	<i>internet</i>
49	teknik proteksi : <i>prefentif</i> pribadi sebelum penetrasi terjadi	<i>computer security</i>	<i>computer security</i>	61	3 2 1 proteksi gratis saat anak ber-internet	<i>internet</i>	<i>internet</i>
50	kriptografi keamanan internet dan jaringan komunikasi	<i>computer security</i>	<i>computer network</i>	62	<i>coreldraw</i> x4	<i>computer graphics</i>	<i>computer graphics</i>
51	memahami model enkripsi & <i>security data</i>	<i>computer security</i>	<i>computer security</i>	63	kreasi desain flora dengan <i>coreldraw</i> x3	<i>computer graphics</i>	<i>computer graphics</i>
52	<i>hacker</i> & keamanan	<i>computer security</i>	<i>microcomputers</i>	64	desain pop art untuk orang awam	<i>computer graphics</i>	<i>computer graphics</i>
53	ancaman internet hacking dan trik menanganinya	<i>computer security</i>	<i>computer security</i>	65	115 teknik profesional <i>corel draw</i> x4	<i>computer graphics</i>	<i>computer graphics</i>
				66	110 trik rahasia <i>corel</i> x4 : mengungkap dan menggali rahasia tersembunyi <i>software</i> grafik populer <i>corel</i> x4	<i>computer graphics</i>	<i>computer graphics</i>

No	Judul	Klasifikasi oleh Perpustakaan Undika	Kategori sistem
67	data nomor pokok perpustakaan (npp) perpustakaan khusus dan sekolah se jawa timur tahun 2015	library	library
68	matabaru penelitian perpustakaan dari serqual ke libqual	library	library
69	ragam inspirasi perpustakaan rumah	library	library
70	dasar - dasar operations research, edisi 2	operations research	operations research
71	teori dan soal-soal operations research (schaum series)	operations research	operations research
72	microprocessor databook	microprocessors	microprocessors
73	dari chip ke sistem pengantar mikroprosesor	microprocessors	microprocessors
74	artificial intelligence	artificial intelligence	data communication
75	mengenal artificial intelligence	Artificial Intelligence	computer science
76	analisa dan perancangan sistem pengolahan data komunikasi data	system analysis	system analysis
77	komunikasi data	data communication	data communication
78	pengolah mikro/komputer mikro (seri pelajaran elektronika)	MicroComputers	computers

Evaluasi

Berikut hasil evaluasi yang didapatkan dari implementasi perancangan sistem klasifikasi kategori. Untuk menentukan Akurasi klasifikasi akan menggunakan rumus berikut :

$$Akurasi = \frac{\text{Jumlah Kategori Benar}}{\text{Jumlah dokumen uji}} \times 100\%$$

$$Akurasi = \frac{62}{78} \times 100\%$$

$$Akurasi = 79.48\% = 79\%$$

Berdasarkan Pengujian Akurasi diatas maka, didapatkan hasil kategori yang sesuai adalah sejumlah 62 dan kategori yang tidak sesuai sejumlah 16 dan hasil akurasi klasifikasi kategori menggunakan algoritma NBC sebesar 79%. Hal ini menunjukkan bahwa algoritma NBC dapat digunakan sebagai algoritma untuk pengklasifikasian kategori buku karena tingkat akurasinya yang cukup tinggi. Persentase uji akurasi akan semakin tinggi jika penggunaan data *Training* semakin banyak dan uji akurasi yang didapatkan diatas berdasarkan data *Training* sejumlah 311 data dengan lama waktu untuk menyiapkannya yaitu 10 menit 11 detik. Dikarenakan pada penelitian ini menggunakan perbandingan data pada kamus kategori maka banyaknya data pada kamus kategori juga dapat mempengaruhi nilai akurasi dari hasil pengujian. Lama waktu klasifikasi kategori dengan sistem hanya membutuhkan 2 detik 70 milidetik yang mana sebelumnya ketika dilakukan secara manual klasifikasi kategori dapat memakan waktu 3 sampai 10 menit. Hal ini menunjukkan bahwa klasifikasi kategori dengan sistem klasifikasi kategori menggunakan algoritma NBC dapat dilakukan menjadi lebih cepat dan efektif. Sistem yang sudah dirancang telah didemokan kepada pihak Perpustakaan Undika dan menghasilkan beberapa saran perbaikan dan pengembangan. Perbaikan yang sudah dilakukan yaitu menampilkan nomor DDC dan penggunaan font yang informative. Pengembangan sistem disarankan untuk membuat inputan kategori baru agar lebih fleksibel digunakan. Sesuai hasil dari google form pada lampiran 5 hasil *Google Form*. Tingkat kepuasan dari penerimaan sistem menunjukkan sesuai/baik yaitu 43%, cukup sesuai/baik yaitu 43%, dan belum sesuai/baik yaitu 14%. Dari persentase menunjukkan hasilnya yaitu 43% sesuai atau cukup sesuai.

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, sistem ini dapat melakukan pengklasifikasian buku dengan memasukan judul atau deskripsi buku berdasarkan DDC 23 pada kelas 000 yaitu *Computer science, information & general works* diproses menggunakan tahapan *text preprocessing* dengan perbandingan kamus kategori dan algoritma NBC.

1. Hasil pengujian akurasi yaitu sebesar 79% yang menunjukkan bahwa kesesuaian hasil klasifikasinya cukup tinggi.
2. Klasifikasi kategori dilakukan dalam waktu yang lebih cepat dengan sistem hanya membutuhkan 2 detik 70 milidetik.

RUJUKAN

- Ginting, S. L., & Trinanda, R. P. (2013). Penggunaan Metode Naïve Bayes Classifier Pada Aplikasi Perpustakaan. *Jurna Teknik Komputer Unikom*, 1-6.
- Indriyono, B. V., Utami, E., & Sunyoto, A. (2015). Klasifikasi Jenis Buku Berdasarkan Judul dan Sinopsis Menggunakan Metode Naïve Bayes Classifier (Studi kasus : STMIK Kadiri). 476-485.
- Librian, A. (2017, February 21). *Sastrawi*. Retrieved September 6, 2020, from Github:
<https://github.com/sastrawi/sastrawi/blob/master/README.en.md>
- Pressman, R. S., & Maxim, B. R. (2015). Software Engineering: A Practitioner's Approach. In R. S. Pressman, & B. R. Maxim, *Software Engineering: A Practitioner's Approach* (pp. 41-43). New York: McGraw-Hill Education.
- Rizqiyani, v., Mulwida, A., & Putri, R. D. (2017). Klasifikasi Judul Buku dengan Algoritma Naive Bayes dan Pencarian Buku pada Perpustakaan Jurusan Teknik Elektro. 60-65.
- Setiawan, A., Astuti, I. F., & Kridalaksana, A. H. (2015). Klasifikasi dan Pencarian Buku Referensi Akademik Menggunakan Metode Naive Bayes

Classifier (NBC) (Studi Kasus: Perpustakaan Daerah Provinsi Kalimantan Timur). *Jurnal Informatika Mulawarman*, 1-10.