

## OPTIMALISASI PROSES PENCARIAN BUKU DONGENG BERBASIS WEB DENGAN MENERAPKAN METODE KLASIFIKASI NAÏVE BAYES

Maria Sri Wulandari, Rahayu Noveandini dan Fajar Nugroho  
STMIK Jakarta STI&K  
Jl. BRI No.17, Radio Dalam, Kebayoran Baru, Jakarta Selatan 12140  
mswuland@yahoo.com, {ayu.noveandini, samafugu}@gmail.com

### ABSTRAK

*Pengklasifikasian dan pencarian buku adalah rangkaian proses pencarian data-data tentang buku dengan tujuan untuk memperoleh sebuah data, informasi tentang buku yang diinginkan. Sistem pencarian sebuah data dapat dilakukan melalui sebuah media digital yang disebut dengan aplikasi. Penelitian ini akan membangun aplikasi yang digunakan untuk pengembangan sistem pencarian buku secara otomatis dengan menerapkan metode klasifikasi Naive Bayes kedalam aplikasi berbasis website. Peran aplikasi pada server sebagai proses untuk memasukan data-data buku yang dilakukan oleh admin dan peran aplikasi pada client adalah sebagai media digital untuk mencari data-data buku oleh pengguna. Sistem pencarian data buku dongeng dengan metode Naive Bayes Clasifier sedangkan proses dalam pengembangan sistem menggunakan tahapan System Development Life Cycle dengan menggunakan model Waterfall. Sistem ini dibangun dengan menggunakan Framework PHP yaitu CodeIgniter dan database MySQL. Model pengembangan perangkat lunak menggunakan Unified Modeling Language yaitu Use Case Diagram, Activity Diagram, Diagram Sequence dan Class Diagram. Penerapan metode klasifikasi Naive Bayes didalam penelitian ini sebagai teknik pencarian terhadap data buku dengan menggunakan kata kunci atau kategori pencarian dari judul buku, kategori buku, dan pengarang buku yang diinput oleh pengguna, jika buku yang dicari tersedia maka sistem akan memperlihatkan data-data buku berdasarkan dari kata-kata kunci atau kategori pencarian yang diinput. Hasil evaluasi dari pengujian akurasi diperoleh sebesar 52,17% yang menunjukkan bahwa algoritma klasifikasi naive bayes dapat diterapkan sebagai algoritma untuk pencarian data buku dongeng.*

**Kata Kunci :** *Sistem Pencarian, Buku Dongeng, Website, Naive Bayes Classifier*

### PENDAHULUAN

Saat ini, perkembangan teknologi informasi semakin meningkat. Perkembangan teknologi ini menyebabkan perusahaan dan organisasi mampu mengumpulkan dan mendapatkan data dari berbagai sumber dengan lebih mudah dan lebih cepat, yang mengakibatkan jumlah data menjadi semakin besar dan akan terus bertambah (Abdi Pandu Kusuma dkk, 2016). Semakin besarnya jumlah data akan berdampak adanya masalah dalam pengklasifikasiannya. Data yang tersebar tanpa diklasifikasikan dengan kategori tertentu, akan mengakibatkan proses pencarian menjadi lambat. Untuk data dengan skala kecil, hal ini tidak akan menimbulkan masalah yang signifikan terhadap proses pencarian data, namun sangat berpengaruh jika data dalam skala yang besar, dimana akan sangat diperlukan sebuah proses pencarian data yang lebih

cepat, sehingga tahap awal perlu dilakukan pengelompokan data (Yakub, 2012).

Naïve Bayes adalah salah satu metode klasifikasi untuk pengelompokan atau pengklasifikasian data mining. Proses kinerja dari metode klasifikasi Naïve Bayes dengan menggunakan formula perhitungan probabilitas. Konsep dasar metode klasifikasi naïve bayes yaitu dengan Teorema Bayes, merupakan teorema yang diterapkan pada statistika untuk proses perhitungan terhadap peluang (Agus Setiawan, 2015). Beberapa penelitian dengan menggunakan naïve bayes yang sudah dilakukan, diantaranya diantaranya Abdi Pandu Kusuma, Ida Sri Rahayu, Sistem Pencarian Katalog Buku Menggunakan Metode Klasifikasi Naïve Bayes dengan penerapan metode klasifikasi Naïve Bayes untuk model pencarian data-data buku dengan kategori judul buku yang dijadikan sebagai kata kunci (Abdi Pandu Kusuma

dkk,2016). Agus Setiawan, Indah Fitri Astuti dan Awang Harsa Kridalaksana, Klasifikasi dan pencarian buku referensi akademik menggunakan metode naïve bayes classifier (NBC) (Studi kasus : Perpustakaan daerah Provinsi Kalimantan Timur), Februari 2015, penelitian yang dilakukan adalah membangun aplikasi untuk pencarian buku (Agus Setiawan dkk, 2015). Faisal, A. Muhammad Syafar, Ummi Azizah Mukaddim, Rekayasa Klasifikasi Pencarian Abstrak Tentang Mikrokontroler e-Journal Instek dengan Algoritma Naive Bayes, Juni 2020, (Faisal dkk, 2020). Guruh Agung Prasetyo, M. Taufik, ST. MIT, Sam Farisa c, ST. M. Kom, Aplikasi Berbasis Web Untuk Deteksi Bahasa Daerah Di Indonesia Dengan Naive Bayes, hasil yang diperoleh dalam deteksi dalam pengembangan aplikasi adalah menjembatani masyarakat yang tidak paham dialek daerah tertentu (Guruh Agung Prasetyo dkk, 2019). Muthia Anggraini, Rizki Ayuning Tyas, Ismi Ana Sulasiyah dan Qurrotul Aini, Implementasi Algoritma Naive bayes Dalam Penentuan Rating Buku, September 2020, tujuan aplikasi ini adalah membuat rating sebuah buku untuk acuan pembaca dengan akurasi 66.6%. [Muthia Anggraini dkk, 2020]. Deri Lianda dan Niko Surya Atmaja, Prediksi data buku favorit menggunakan metode naïve bayes, Februari 2021, penelitian ini menerapkan data mining pencarian data buku (Deri Lianda dkk, 2021).

Berdasarkan artikel penelitian yang sudah dilakukan sebelumnya, dengan menerapkan metode naïve bayes untuk sistem pencarian dan klasifikasi, maka didalam penelitian ini kami akan membangun sebuah sistem pencarian data buku dongeng berbasis web, sistem ini berfokus menyediakan buku dongeng untuk anak-anak dengan kumpulan cerita yang menarik. Tujuan dari pembuatan sistem informasi buku digital dongeng anak ini adalah 1) Terbentuknya media digital yang dapat memberikan kemudahan bagi anak-anak dalam mengakses buku tentang dongeng, 2) Terbentuknya sistem informasi buku digital tentang dongeng anak-anak berbasis web, 3) Penerapan metode Naive Bayes Classifier untuk melakukan pencarian buku.

Sistem ini diharapkan dapat memudahkan bagi anak-anak untuk mencari dan menemukan buku dongeng yang diinginkan. Sistem ini diharapkan sangat membantu bagi anak-anak dengan kondisi saat ini, dimana segala aktifitas sebisa mungkin dilakukan secara on line. Namun dengan banyaknya buku dongeng akan menyulitkan anak-anak dalam mencari buku yang mereka inginkan. Oleh karena itu sistem pencarian data buku dongeng berbasis web ini menggunakan metode Naive Bayes Classifier (NBC) dengan menerapkan algoritma probabilistic learning method, yaitu pada saat klasifikasi dengan mencari probabilitas tertinggi berdasarkan atribut yang diinput (Yakub, 2012).

## METODE NAÏVE BAYES

Penelitian ini akan dilakukan dengan memanfaatkan metode klasifikasi Naive Bayes Classifier sebagai sistem pencarian buku dongeng. Metode klasifikasi Naive Bayes dapat melakukan suatu prediksi terhadap probabilitas kedepannya atas dasar pengalaman yang terjadi pada masa terdahulunya. Ciri khas metode Naive Bayes Classifier memiliki asumsi kuat terhadap independensi dari setiap dan semua kejadian (Abdi Pandu Kusuma dkk, 2016).

Formula dasar Naive Bayes yang digunakan untuk pemrograman adalah: (Rifki Mifathur Sutomo dkk, 2020)

$$P(M|N) = P(M|N) * P(M) / P(N) \quad (1)$$

Peluang suatu kejadian M sebagai N diperoleh dari suatu peluang N saat M, peluang M dan peluang N. Maka terjadi perubahan formula sebagai berikut :

$$P(S_i|T) = P(T|S_i) * P(S_i) / P(T) \quad (2)$$

Metode Klasifikasi Naive Bayes adalah model penyederhanaan dari Metode Bayes yang sesuai dalam proses klasifikasi teks atau dokumen. Persamaan yang terbentuk sebagai berikut:

$$V_{MAP} = \arg \max P(V_j | X_1, X_2, \dots, X_n) \quad (3)$$

Berdasarkan persamaan ke 3, akan terbentuk persamaan ke 4:

$$V_{MAP} = \arg \max_{v_j \in V} \frac{P(x_1, x_2, \dots, x_n | v_j) P(v_j)}{P(x_1, x_2, \dots, x_n)} \quad (4)$$

$P(x_1, x_2, \dots, x_n)$  adalah konstan, maka dapat berubah menjadi :

$$V_{MAP} = \arg \max_{v_j \in V} P(x_1, x_2, \dots, x_n | v_j) P(v_j) \quad (5)$$

$P(x_1, x_2, \dots, x_n | v_j)$  kesulitan untuk dilakukan perhitungan, sehingga dapat dinyatakan bahwa setiap kata yang terdapat didalam dokumen, ternyata tidak ada keterkaitan.

$$V_{MAP} = \arg \max_{v_j \in V} P_{v_j} \prod_i P(v_i | v_j) \quad (6)$$

Penjelasan :

$$P(v_i) = \frac{|frek|}{|jumlah|} \quad (7)$$

$$P(Y_k | v_j) = \frac{b_k + 1}{b + |kons tan ta|} \quad (8)$$

Keterangan

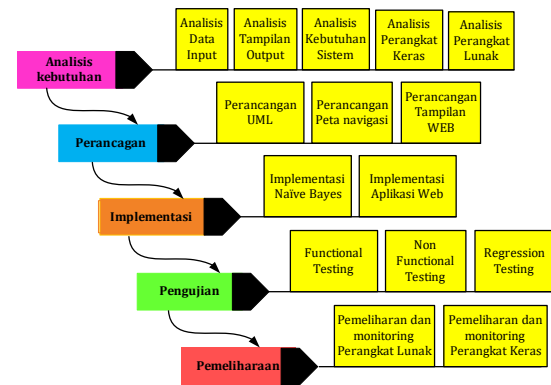
1.  $P(v_j)$  : Probabilitas untuk setiap dokumen pada sekelompok dokumen
2.  $P(Y_k | v_j)$  : Probabilitas kemunculan terhadap kata “ $Y_k$ ” untuk dokumen dengan judul class  $v_j$
3.  $|frek|$  : frekuensi dokumen untuk tiap judul
4.  $|jumlah|$  : jumlah dokumen yang tersedia.
5.  $b_k$  : frekuensi sebuah kata ke-k untuk tiap judul Kosakata: jumlah kata didalam dokumen test

Untuk persamaan (8) dilakukan proses penambahan 1 terhadap pembilang, dengan maksud untuk mengantisipasi apabila muncul suatu kata didalam dokumen uji yang tidak terdapat pada setiap

dokumen data latih (Agus Setiawan dkk,2015).

## METODE PENELITIAN

Tahapan pengembangan sistem menggunakan SDLC dengan model waterfall, yang meliputi tahapan sebagai berikut Analisis Kebutuhan, Perancangan, Implementasi, Pengujian dan Pemeliharaan. Seperti terlihat pada Gambar 1 dibawah ini (Sutanta dkk,2011).



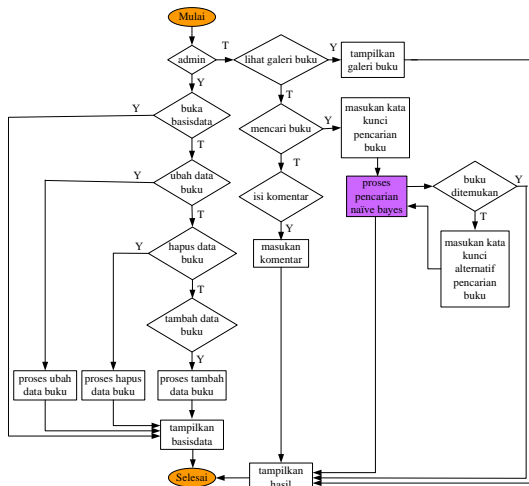
Gambar 1. Tahapan Pengembangan Sistem

## HASIL DAN PEMBAHASAN

Analisa masalah yang terdapat pada penelitian ini adalah anak-anak mengalami kesulitan dalam mendapatkan informasi dan mengakses buku tentang dongeng. Dengan adanya permasalahan ini maka akan dibangun sistem pencarian buku tentang dongeng berbasis web.

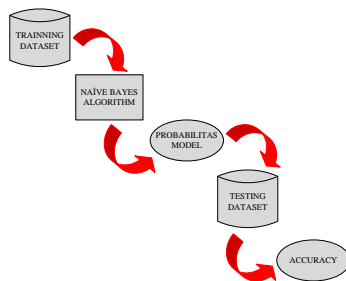
Sistem yang dibangun diharapkan dapat membantu mengurangi kesulitan dalam pencarian informasi tentang buku dongeng anak-anak. Metode Naive Bayes Classifier adalah metode klasifikasi yang dapat digunakan untuk menentukan probabilitas berdasarkan kata kunci yang di cari.

Gambar 2 menggambarkan diagram alir alir perangkat lunak sistem pencarian buku dongeng dari sudut pandang admin dan pengguna. Dari bagian admin, admin dapat melakukan proses menambah, mengubah dan menghapus data buku; sedangkan sebagai pengguna, diberikan fasilitas melihat galeri buku, memberikan komentar dan mencari buku berdasarkan kata kunci judul, kategori dan penerbit.



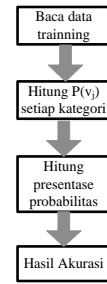
**Gambar 2** Diagram alir perangkat lunak sistem pencarian buku dongeng

Tahapan pada proses klasifikasi ditampilkan pada gambar 2. Dataset dibagi menjadi dua, untuk 80% data set untuk proses training dan 20% dataset untuk proses testing. Training dataset dilakukan dengan menggunakan algoritma naive bayes sehingga akan menghasilkan sebuah model. Menggunakan model tersebut akan dilakukan testing dataset untuk menghasilkan tingkat akurasi yang optimal.



**Gambar 3** Tahapan Klasifikasi

Proses naive bayes diuraikan pada gambar 4, dengan menggunakan data training akan dilakukan perhitungan  $P(v_j)$  untuk setiap kategori buku dongeng. Kemudian dihitung presentase probabilitas untuk menghasilkan tingkat akurasi.



**Gambar 4** Tahapan Naive Bayes

Pengujian dari metode Naive Bayes Classifier dengan menggunakan kata kunci untuk mencari buku berdasarkan judul dan dapat menghasilkan data yang akurat :

$$P(v_j) = \frac{|docs_j|}{|Contoh|}$$

Keterangan :

$|docs_j|$  : Frekuensi sebuah buku untuk tiap kategori (Kemunculan dari setiap kata menurut kata kunci)

$|Contoh|$  : Total buku yang ada

Pengujian ini dilakukan dengan cara memasukan kata kunci “kisah” yang dimasukan oleh user dengan menggunakan metode Naive Bayes Classifier. Data dari pencarian dengan menggunakan kata kunci “kisah” dapat dilihat pada Tabel 1

**Tabel 1** Pengujian Metode Naive Bayer Classifier

No.	docs	Contoh	P(Vj)	Hasil
1.	4	23	4/23	0,1739130434782609
2.	12	23	12/23	0,5217391304347826
3.	1	23	1/23	0,0434782608695652

Melihat dari nilai  $P(v_j)$  akan dipilih nilai tertinggi atau nilai yang terbesar untuk dimasukan pada tabel 2 Probabilitas Klasifikasi Naive Bayes akan ditampilkan sebagai nilai dari hasil akurasi proses pencarian didalam aplikasi. Kode kategori buku digunakan untuk membedakan setiap kategori.

Keterangan:

1. KCR1, Kategori Cerita Rakyat.
2. KFB1, Kategori Fabel.
3. KLG1, Kategori Legenda.

**Tabel 2** Probabilitas Naive Bayer Classifier

No.	Kode Kategori Buku	Naive Bayes
1.	KCR1	0,1739130434782609
2.	KFB1	0,5217391304347826
3.	KLGI	0,0434782608695652

Probabilitas merupakan pengukuran suatu kemungkinan sebuah peristiwa yang akan terjadi. Nilai probabilitas dinyatakan dengan angka 0 sampai 1 dan dapat dinyatakan dalam bentuk desimal 1=100%. Nilai Probabilitas = 0 menjelaskan bahwa sebagai peristiwa yang tidak mungkin terjadi. Nilai Probabilitas = 1 merupakan sebuah peristiwa yang pasti terjadi. Sehingga definisi probabilitas adalah sebagai peluang suatu kejadian, seperti yang terlihat pada tabel 3

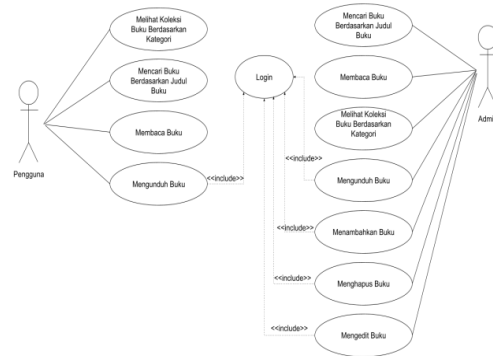
$$\text{Persentase } P(v_j) = \frac{\text{Nilai } P(v)}{P(v)_{\max}} \times 100\%$$

**Tabel 3** Persentase Probabilitas

No.	Kode Kategori Buku	Naive Bayes	Persentase P(Vj)
1.	KCR1	0,1739130434782609	17.39%
2.	KFB1	0,5217391304347826	52.17%
3.	KLGI	0,0434782608695652	4.35%

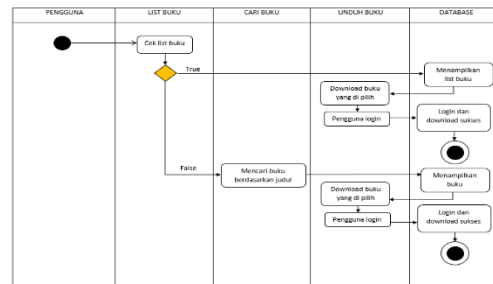
Tabel 3 menjelaskan sebuah persentase probabilitas yang menghasilkan kesimpulan bahwa proses pencarian buku dengan menggunakan kata kunci “Kisah” dengan metode Naive Bayes Classifier pada sistem pencarian buku dongeng anak berbasis web terdapat akurasi tertinggi 52.17%.

Pengembangan perangkat lunak dengan menggunakan Unified Modeling Language (UML). UML merupakan alat bantu sering disebut dengan tools biasa digunakan pada bahasa pemrograman berorientasi objek. Berikut adalah rangkaian diagram UML (Sutanta dkk,2011).



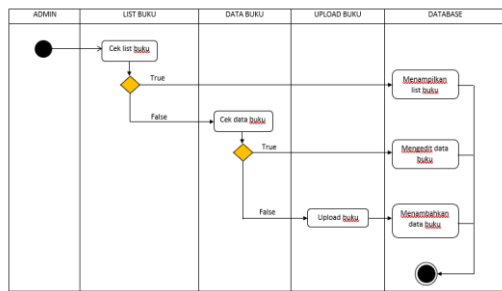
**Gambar 5.** Use Case Diagram

Gambar 5 menggambarkan admin dan pengguna dalam menggunakan sistem dan bagaimana perilaku admin serta pengguna terhadap sistem.



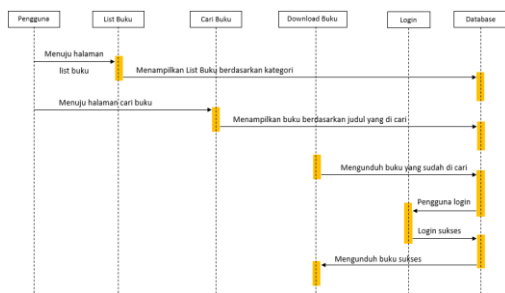
**Gambar 6.** Diagram Activity Pengguna

Di jelaskan pada gambar 6 yang menjadi actor pada activity diagram adalah pengguna. Pada alur activity diagram pengguna yang pertama adalah apakah pengguna mau melihat list buku yang ada di website. Apabila pengguna ingin melihat list buku maka database akan menampilkan seluruh list buku yang ada, pengguna dapat mengunduh salah satu buku tersebut ketika pengguna sudah melakukan login. Atau pengguna dapat mencari buku sesuai dengan judul buku, pengguna dapat mengunduh buku tersebut ketika pengguna sudah melakukan login .



Gambar 7. Diagram Activity Admin

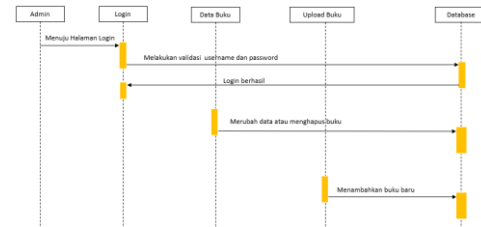
Pada gambar 7 yang menjadi actor pada activity diagram adalah admin. Berbeda dengan alur activity diagram pengguna, pada activity diagram admin mempunyai alurnya tersendiri. Ketika admin mengakses website yang pertama kali admin lakukan adalah memilih menu mana yang akan di akses. Menu yang pertama list buku, admin dapat melihat semua koleksi buku yang ada di website. Menu selanjutnya data buku, admin dapat mengelola data buku seperti merubah nama buku, penulis, penerbit, tahun terbit dan sinopsis buku yang sudah ada di website. Terakhir upload buku, admin dapat menambahkan koleksi buku baru di halaman ini.



Gambar 8. Diagram sequence Pengguna

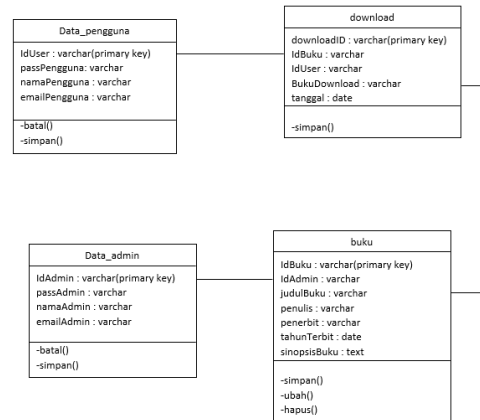
Pada gambar 8 merupakan penjelasan dari urutan penggunaan sistem oleh pengguna. Pengguna sistem digambarkan sebagai aktor yang melakukan tahapan berinteraksi dengan antarmuka sistem. Pada diagram sequence Pengguna dapat mengakses fitur list buku dan cari buku tanpa harus melakukan login terlebih dahulu. Fitur list buku berfungsi untuk menampilkan buku berdasarkan kategori buku dan fitur cari buku berfungsi untuk mencari buku berdasarkan judul buku

yang di cari. Selain itu pengguna dapat melakukan download / mengunduh buku yang sudah di cari namun pengguna diharuskan untuk login terlebih dahulu.



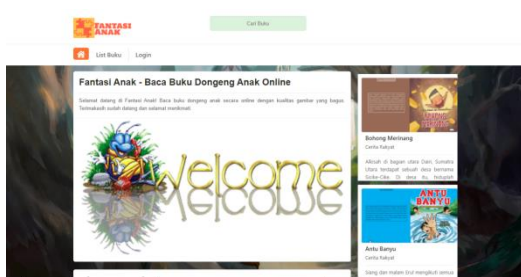
Gambar 9 Diagram sequence admin

Pada gambar 9 yang menjadi Actor adalah Admin. Pada diagram sequence admin sebelum melakukan aktifitas admin diwajibkan untuk login melalui halaman masuk agar dapat mengakses halaman berikutnya. Setelah berhasil admin dapat melakukan aktifitas seperti mengelola data buku seperti mengedit judul buku, nama penulis, tahun terbit dan sinopsis buku atau menghapus buku dari koleksi buku. Upload buku digunakan untuk menambahkan buku baru ke dalam koleksi.



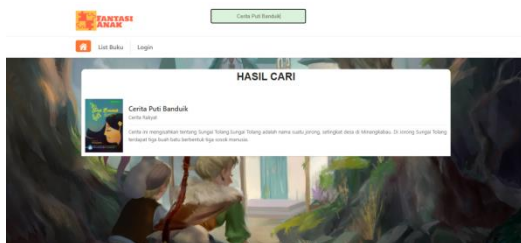
Gambar 10. Class diagram

Gambar 10 adalah class diagram yang menggambarkan objek yang memiliki atribut dan method yang pada sistem. Setiap objek menggambarkan data yang terdapat pada sistem. Setiap objek memiliki relasi dengan objek lain sesuai relasi data pada sistem. Terdapat 4 class yaitu Data\_pengguna, Data\_admin, download dan buku.



Gambar 11 Tampilan Halaman Beranda

Pada gambar 11 halaman beranda adalah halaman awal ketika pengguna dan admin mengakses website. Halaman ini menampilkan beberapa koleksi buku dongeng yang di rekomendasikan oleh admin.



Gambar 12 Tampilan Halaman Cari Buku

Pada gambar 12 halaman cari buku adalah halaman yang nantinya akan digunakan untuk mencari koleksi buku dongeng yang ada berdasarkan dari judul yang di cari.

Pengujian perangkat lunak dengan menggunakan teknik blackbox. Beberapa pengujian yang dilakukan diantaranya pengujian terhadap fitur-fitur yang disajikan pada website, dengan menguji fungsi fitur secara keseluruhan. Pengujian berikutnya adalah untuk menguji sistem jika diimplementasikan pada perangkat, sistem operasi dan browser yang berbeda. Sistem juga dilakukan pengujian apabila dilakukan upgrade software, untuk melihat sistem tetap stabil atau terjadi regresi, misalkan sistem tidak bisa dijalankan pada versi yang lebih tinggi.

## PENUTUP

Berdasarkan hasil penelitian dengan menerapkan metode klasifikasi naïve bayes pada sistem pencarian buku dongeng anak anak, penelitian ini setelah dilakukan simulasi, dapat ditarik kesimpulan sebagai berikut:

1. Sistem pencarian buku dongeng berbasis web yang dibangun melalui penelitian ini dapat dimanfaatkan sebagai solusi alternatif untuk meminimalkan masalah dalam pencarian buku dongeng bagi anak-anak.
2. Penerapan metode klasifikasi Naive Bayes dapat dilakukan dengan menggunakan kata kunci berdasarkan judul buku, kategori buku, dan pengarang buku.
3. Hasil yang diperoleh dari penelitian ini menunjukkan bahwa pengujian dengan menggunakan metode klasifikasi Naive Bayes menghasilkan presentase keakuratan pada pencarian data dengan presentase probabilitas klasifikasi Naive Bayes sebesar 52.17%.

## DAFTAR PUSTAKA

- [1] Abdi Pandu Kusuma, Ida Srirahayu, "Sistem Pencarian Katalog Buku Menggunakan Metode Naïve Bayes Classifier (NBC) Pada Aplikasi Mulia-Bookstore Berbasis Android", Jurnal Antivirus, Vol. 10 No. 2 November 2016, p-ISSN: 1978-5232 e-ISSN: 2527-337X.
- [2] Agus Setiawan, Indah Fitri Astuti dan Awang Harsa Kridalaksana, "Klasifikasi dan pencarian buku referensi akademik menggunakan metode naïve bayes classifier (NBC)(Studi kasus : Perpustakaan daerah Provinsi Kalimantan Timur)", Jurnal Informatika Mulawarman, Vol. 10 No. 1, Februari 2015.
- [3] Deri Lianda, Niko Surya Atmaja, "Prediksi Data Buku Favorit Menggunakan Metode Naïve Bayes", Jurnal Pseudocode, Volume VIII Nomor 1, Februari 2021.
- [4] Faisal, A. Muhammad Syafar, Umami Azizah Mukaddim, "Rekayasa Klasifikasi Pencarian Abstrak Tentang Mikrokontroler e-Journal Instek dengan Algoritma Naive Bayes", JAMI Jurnal Ahli Muda

- Indonesia Vol. 1 No. 1 (2020): Juni 2020.
- [5] Guruh Agung Prasetyo, M.Taufik, ST.MIT, Sam Farisa c,ST.M.Kom, “Aplikasi Berbasis Web Untuk Deteksi Bahasa Daerah Di Indonesia Dengan Naïve Bayes”, Konferensi Ilmiah Mahasiswa UNISSULA (KIMU) Universitas Islam Sultan Agung Semarang, 18 Oktober 2019.
- [6] Muthia Anggraini, Rizki Ayuning Tyas, Ismi Ana Sulasiyah, Qurrotul Aini, “Implementasi Algoritma Naïve bayes Dalam Penentuan Rating Buku”, SISTEMASI: Jurnal Sistem Informasi Volume 9, Nomor 3, September 2020.
- [7] Rifki Mifathur Sutomo, Arie Ardianti Suryani, “Implementasi Algoritma Naïve Bayes untuk Word Sense Disambiguation dalam Bahasa Indonesia”, e-Proceeding of Engineering : Vol.7, No.2 Agustus 2020.
- [8] Sutanta,Edhy. ”Basis Data dalam Tinjauan Konseptual”, Yogyakarta: Andi, 2011.
- [9] Yakub, “Pengantar Sistem Informasi”, Yogyakarta: Graha Ilmu, 2012.