

# Aplikasi Pengarsipan Dokumen Menggunakan Metode String Matching (Studi Kasus Perpustakaan SMP Negeri 5 Seluma)

Muntahanah<sup>1</sup>, Yulia Darnita,<sup>2</sup>

<sup>1</sup>Program Studi Informatika, Fakultas Teknik, Universitas Muhammadiyah Bengkulu

<sup>2</sup>Program Studi Informatika, Fakultas Teknik, Universitas Muhammadiyah Bengkulu

Jl. Bali PO BOX 118. Telp (0736) 227665, Fax (0736) 26161, Bengkulu 38119

E-Mail : muntahanah@umb.ac.id<sup>1</sup>, yuliadarnita@umb.ac.id<sup>2</sup>

*Abstack: The library is to support, facilitate and enhance the quality of the implementation of school activities programs to obtain information or knowledge through information services covering five aspects, namely information gathering, information preservation, information processing, information utilization and information dissemination. In the search process String Matching is very suitable for using the Knuth-Morris-Pratt algorithm which is an algorithm to search for a string consisting of several characters (usually called a pattern), Knuth-Morris-Pratt Algorithm or commonly abbreviated as KMP algorithm to compare text characters and character pattern on the pattern from left to right. The idea of this algorithm is how to utilize the known pattern characters in the text until there is a mismatch to shift. Based on the results of the system testing, it can be concluded: Provide information on the distribution of book data in SMP Negeri 5 Seluma Bengkulu City using string matching method, so that library staff do not need to re-record book data into a new master book if there is a lost and damaged book, and make it easier for students and students to search for books to be read or borrowed. Provide convenience for SMP Negeri 5 Seluma, in accessing book data information based on subjects, and make it easier for students and students to find book information in the library of SMP Negeri 5 Seluma.*

**Abstrak** : Perpustakaan adalah untuk mendukung, memperlancar serta mempertinggi kualitas pelaksanaan program kegiatan Sekolah-sekolah untuk memperoleh informasi atau pengetahuan melalui pelayanan informasi yang meliputi lima aspek yaitu pengumpulan informasi, pelestarian informasi, pengolahan informasi, pemanfaatan informasi dan penyebarluasan informasi. Dalam proses pencarian String Matching sangat cocok untuk menggunakan algoritma Knuth-Morris-Pratt yang merupakan algoritma untuk melakukan pencarian sebuah string yang terdiri dari beberapa karakter (biasa disebut pattern), Algoritma Knuth-Morris-Pratt atau biasa disingkat algoritma KMP melakukan perbandingan karakter-teks dan karakter pattern pada pola dari kiri ke kanan. Ide dari algoritma ini adalah bagaimana memanfaatkan karakter-karakter pattern yang sudah diketahui ada di dalam teks sampai terjadinya ketidakcocokkan untuk melakukan pergeseran. Berdasarkan hasil pengujian sistem yang dilakukan dapat diambil kesimpulan : Memberikan informasi distribusi data buku pada SMP Negeri 5 Seluma Kota Bengkulu dengan metode *string matching*, sehingga petugas perpustakaan tidak perlu mencatat ulang data buku kedalam buku induk yang baru apabila ada buku yang hilang dan rusak, serta memudahkan siswa dan siswi untuk mencari buku yang akan dibaca atau dipinjam, Memberikan kemudahan bagi SMP Negeri 5 Seluma, dalam mengakses informasi data buku berdasarkan mata pelajaran, serta mempermudah siswa dan siswi dalam mencari informasi buku di perpustakaan SMP Negeri 5 Seluma.

## I. PENDAHULUAN

Perpustakaan adalah untuk mendukung, memperlancar serta mempertinggi kualitas pelaksanaan program kegiatan Sekolah-sekolah untuk memperoleh informasi atau pengetahuan melalui pelayanan informasi yang meliputi lima aspek yaitu pengumpulan informasi, pelestarian informasi, pengolahan informasi, pemanfaatan informasi dan penyebarluasan informasi.

Peningkatan kinerja dan pelayanan, penggunaan dan pemanfaatan teknologi informasi semakin diperlukan untuk menggantikan peran sistem berbasis komputer yang memiliki kemampuan dalam menangani data bereferensi geografi yaitu pemasukan data, manajemen data (penyimpanan dan pemanggilan kembali), memanipulasi dan analisis data, serta keluaran sebagai hasil akhir (output). Kemudahan yang dapat diperoleh dari penggunaan teknologi informasi untuk pengolahan data maupun penyajian informasi yang cepat dan akurat sangat membantu untuk meningkatkan kinerja organisasi agar lebih efisien dan efektif. Namun, seringkali banyak

sekolah-sekolah yang mengabaikan atau kurang memaksimalkan pemakaian teknologi informasi di dalam kegiatan operasionalnya..

Pencarian kata atau judul Buku sangat perlu mengingat judul buku satu dengan yang lain saling berhubungan atau berkaitan, sehingga dengan menggunakan satu kata kunci atau lebih dari judul buku, dapat dengan cepat mendapatkan tertulis dimana buku tersebut terletak. Dalam proses pencarian String Matching sangat cocok untuk menggunakan algoritma Knuth-Morris-Pratt yang merupakan algoritma untuk melakukan pencarian sebuah string yang terdiri dari beberapa karakter (biasa disebut pattern). Algoritma Knuth-Morris-Pratt merupakan salah satu algoritma string matching dikembangkan secara terpisah oleh James H. Morris bersama Vaughan R. Pratt pada tahun 1966 dan Donald E. Knuth pada tahun 1967, kemudian dipublikasikan secara bersamaan pada tahun 1977. Algoritma Knuth-Morris-Pratt atau biasa disingkat algoritma KMP melakukan perbandingan karakter teks dan karakter pattern pada pola

dari kiri ke kanan. Ide dari algoritma ini adalah bagaimana memanfaatkan karakter-karakter pattern yang sudah diketahui ada di dalam teks sampai terjadinya ketidakcocokkan untuk melakukan pergeseran.

## II. TINJAUAN PUSTAKA

### 2.1. Perpustakaan

Perpustakaan adalah institusi yang menyediakan koleksi bahan pustaka tertulis, tercetak dan terekam sebagai pusat informasi yang diatur menurut sistem aturan dan didayagunakan untuk keperluan pendidikan, penelitian, serta rekreasi intelektual bagi masyarakat. Perpustakaan berperan melakukan layanan informasi literal kepada masyarakat [1].

### 2.2. Dokumen

Ahli dalam dua pengertian, yaitu: 1) Sumber tertulis bagi informasi sejarah sebagai kebalikan dari pada kesaksian lisan, artefak, peninggalan-peninggalan terlukis, dan petilasan-petilasan arkeologis. 2) Diperuntukan bagi surat-surat resmi dan surat-surat negara seperti surat perjanjian, undang-undang, hibah, konsesi, dan lainnya. Lebih lanjut, Gottschalk menyatakan bahwa dokumen (dokumentasi) dalam pengertiannya yang lebih luas berupa setiap proses pembuktian yang didasarkan atas jenis sumber apapun, baik itu yang bersifat tulisan, lisan, gambaran, atau arkeologis. Dari pengertian yang telah dijabarkan oleh Louis Gottschalk dapat disimpulkan bahwa pengertian dari dokumen adalah segala jenis catatan tertulis, gambar atau rekaman yang berkaitan dengan keperluan pengelolaan baik bersifat hardcopy ataupun softcopy [2].

### 2.2 Arsip

Arcief dalam bahasa Belanda mempunyai pengertian sebagai berikut : a. Tempat penyimpanan secara teratur bahan-bahan arsip: bahan-bahan tertulis, piagam-piagam, surat-surat, keputusan-keputusan, akte-akte, daftar-daftar dan dokumendokumen. b. Kumpulan literatur dari bahan-bahan kearsipan tersebut di atas. c. Bahan-bahan yang harus diarsip itu sendiri [3]

Pengertian arsip dalam bidang Administrasi Negara (LAN) adalah naskah-naskah, baik dalam bentuk tunggal maupun dalam bentuk kelompok atau kumpulan, baik dalam bentuk tertulis/bergambar maupun dalam bentuk suara (rekaman) [4].

### 2.3 Algoritma

Algoritma adalah urutan logis langkah-langkah penyelesaian masalah. Masalah tersebut dapat berupa apa saja, dengan catatan untuk setiap masalah, ada kriteria kondisi awal yang harus dipenuhi sebelum menjalankan algoritma. Algoritma akan dapat selalu berakhir untuk semua kondisi awal yang memenuhi kriteria, dalam hal ini

berbeda dengan heuristik. Algoritma sering mempunyai langkah pengulangan (iterasi) atau memerlukan keputusan (logika Boolean dan perbandingan) sampai tugasnya selesai [5].

Algoritma adalah sebuah prosedur komputasi yang mentransformasi sejumlah input menjadi sejumlah output. Sebuah algoritma dikatakan “benar (correct)” jika untuk setiap inputnya menghasilkan output yang benar pula. Secara umum algoritma adalah sekumpulan urutan perintah yang menentukan operasi – operasi tertentu yang diperlukan untuk menyelesaikan suatu masalah ataupun mengerjakan suatu tugas [6].

#### A. Pengertian Algoritma String Matching

Algoritma string matching adalah sebuah algoritma yang digunakan dalam pencocokkan suatu pola kata tertentu terhadap suatu kalimat atau teks panjang. Algoritma string matching sendiri dapat dilakukan dengan beberapa cara tertentu [7]

Pengertian string menurut Dictionary of Algorithms and Data Structures, National Institute of Standards and Technology (NIST) adalah susunan dari karakter-karakter (angka, alfabet atau karakter yang lain) dan biasanya direpresentasikan sebagai struktur data array. String dapat berupa kata, frase, atau kalimat.

Pencocokan string merupakan bagian penting dari sebuah proses pencarian string (string searching) dalam sebuah dokumen. Hasil dari pencarian sebuah string dalam dokumen tergantung dari teknik atau cara pencocokan string yang digunakan. Pencocokan string (string matching) menurut Dictionary of Algorithms and Data Structures, National Institute of Standards and Technology (NIST), diartikan sebagai sebuah permasalahan untuk menemukan pola susunan karakter string di dalam string lain atau bagian dari isi teks.

Pencocokan string (string matching) secara garis besar dapat dibedakan menjadi dua yaitu :

1. Exact string matching, merupakan pencocokan string secara tepat dengan susunan karakter dalam string yang dicocokkan memiliki jumlah maupun urutan karakter dalam string yang sama. Contoh : kata step akan menunjukkan kecocokan hanya dengan kata step.
2. Inexact string matching atau Fuzzy string matching, merupakan pencocokan string secara samar, maksudnya pencocokan string dimana string yang dicocokkan memiliki kemiripan dimana keduanya memiliki susunan karakter yang berbeda (mungkin jumlah atau urutannya) tetapi string-string tersebut memiliki kemiripan baik kemiripan tekstual/penulisan (approximate string matching) atau kemiripan ucapan (phonetic string matching). Inexact string matching masih dapat dibagi menjadi dua yaitu :
  - a. Pencocokan string berdasarkan kemiripan penulisan (approximate string matching) merupakan pencocokan string dengan dasar kemiripan dari segi penulisannya (jumlah karakter, susunan karakter dalam dokumen).

Tingkat kemiripan ditentukan dengan jauh tidaknya beda penulisan dua buah string yang dibandingkan tersebut dan nilai tingkat kemiripan ini ditentukan oleh pemrogram (programmer). Contoh : c mpuler dengan compiler, memiliki jumlah karakter yang sama tetapi ada dua karakter yang berbeda. Jika perbedaan dua karakter ini dapat ditoleransi sebagai sebuah kesalahan penulisan maka dua string tersebut dikatakan cocok.

- b. Pencocokan string berdasarkan kemiripan ucapan (phonetic string matching) merupakan pencocokan string dengan dasar kemiripan dari segi pengucapannya meskipun ada perbedaan penulisan dua string yang dibandingkan tersebut. Contoh yang lain adalah step, dengan steppe, sttep, stepp, stepe.

Pengertian string menurut Dictionary of Algorithms and Data Structures, National Institute of Standards and Technology (NIST) adalah susunan dari karakter-karakter (angka, alfabet atau karakter yang lain) dan biasanya direpresentasikan sebagai struktur data array. String dapat berupa kata, frase, atau kalimat [8].

## B. Macam-macam jenis Algoritma String Matching

### 1. Algoritma Knuth-Morris-Pratt

Algoritma Knuth-Morris-Pratt merupakan pengembangan dari algoritma pencarian string sebelumnya, yaitu algoritma Brute Force. Algoritma Brute-Force merupakan algoritma dasar yang paling sederhana dalam menyelesaikan persoalan pencocokan string yang melakukan pencarian pada setiap posisi di dalam teks antara 0 dan n-m, dimana n adalah panjang teks/banyaknya nama file yang tersimpan di komputer dan m adalah panjang karakter dari suatu pattern (kata yang akan dicari).

Secara sistematis, langkah-langkah yang dilakukan algoritma Knuth-Morris-Pratt pada saat mencocokkan string :

- a. Masukkan Query kata yang akan dicari. Dengan permisalan P=Pattern atau pola susunan kata yang dijadikan sebagai contoh atau pola teks yang akan dicari T=Teks atau judul dokumen
- b. Algoritma Knuth-Morris-Pratt mulai mencocokkan pattern atau pola susunan kata yang dijadikan sebagai contoh pada awal teks.
- c. Dari kiri ke kanan, algoritma ini akan mencocokkan karakter per karakter pattern atau pola susunan kata yang dijadikan sebagai contoh dengan karakter di teks yang bersesuaian, sampai salah satu kondisi berikut dipenuhi :
  1. Karakter di pattern atau pola susunan kata yang dijadikan sebagai contoh dan di teks yang dibandingkan tidak cocok (mismatch)
  2. Semua karakter di pattern atau pola susunan kata yang dijadikan sebagai contoh cocok. Kemudian

algoritma akan memberitahukan penemuan di posisi ini.

- d. Algoritma kemudian menggeser pattern atau pola susunan kata yang dijadikan sebagai contoh berdasarkan tabel next, lalu mengulangi langkah no. 2 sampai pattern atau pola susunan kata yang dijadikan sebagai contoh berada di ujung teks [9].

Algoritma Knuth-Morris-Pratt (KMP) adalah salah satu algoritma pencarian string, dikembangkan secara terpisah oleh Donald E. Knuth pada tahun 1967 dan James H. Morris bersama Vaughan R. Pratt pada tahun 1966, namun keduanya mempublikasikannya secara bersamaan pada tahun 1977 [1]. Jika kita melihat algoritma brute force lebih mendalam, kita mengetahui bahwa dengan mengingat beberapa perbandingan yang dilakukan sebelumnya kita dapat meningkatkan besar pergeseran yang dilakukan. Hal ini akan menghemat perbandingan, yang selanjutnya akan meningkatkan kecepatan pencarian. Perhitungan pergeseran pada algoritma ini adalah sebagai berikut, bila terjadi ketidakcocokan pada saat pattern sejajar dengan teks  $[i..i + n-1]$ , kita bisa menganggap ketidakcocokan pertama terjadi di antara teks  $[i + j]$  dan pattern  $[j]$ , dengan  $0 < j[10]$ .

### 2. Algoritma Rabin-Karp

Algoritma Rabin-Karp merupakan algoritma pencocokan string yang menggunakan fungsi hash sebagai pembanding antara string yang dicari (m) dengan substring pada teks (n). Algoritma Rabin-Karp didasarkan pada fakta jika dua buah string sama maka harga hash value-nya pasti sama [11].

### 3. Brute Force Brute

Force adalah sebuah pendekatan langsung (straight forward) untuk memecahkan suatu masalah, yang biasanya didasarkan pada pernyataan masalah (problem statement) dan definisi konsep yang dilibatkan. Pada dasarnya algoritma Brute Force adalah alur penyelesaian suatu permasalahan dengan cara berpikir yang sederhana dan tidak membutuhkan suatu permikiran yang lama [12].

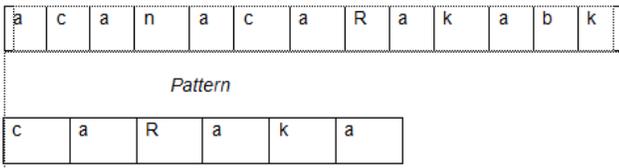
## III. METODE PENELITIAN

### 3.1 Metode Pengumpulan Data

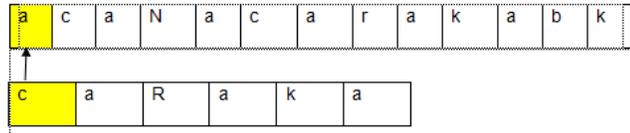
1. Wawancara, suatu cara mengumpulkan data yang digunakan untuk memperoleh informasi langsung dari sumbernya.
2. Observasi, dalam hal ini observasi dilakukan untuk mengamati secara kualitatif berbagai kegiatan dan peristiwa. Dalam penelitian untuk memperoleh data atau informasi yang lebih spesifik kelebihan dan kekurangan pada pencarian buku di perpustakaan SMP Negeri 5 Seluma.
3. Studi pustaka, studi pustaka dilakukan untuk mengetahui secara komprehensif mengenai Algoritma Knuth Morris Pratt mulai dari flowchart yang berhubungan dengan penelitian

3.2 Analisis Algoritma Knuth-Morris-Prat

1. Tentukan Nilai S dan *pattern*  
String S

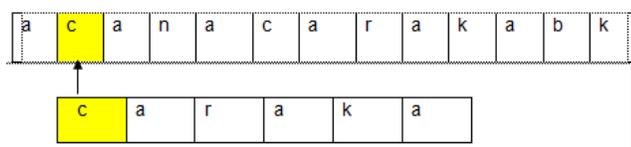


2. Bandingkan *pattern* [1] dan *string* [1]

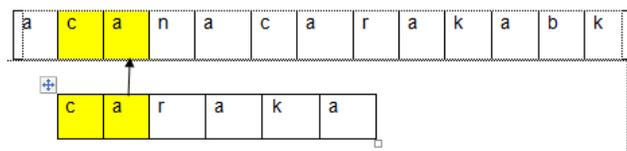


*Pattern* [1] tidak cocok dengan *string* [1]. Maka *pattern* akan bergeser satu Posisi ke kanan.

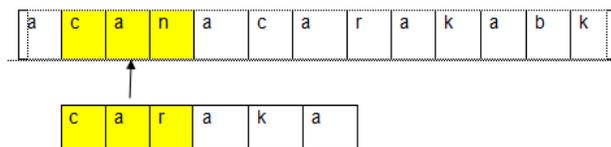
3. Bandingkan *pattern* [1] dan *string* [2]



4. Bandingkan *pattern* [2] dan *string* [3]

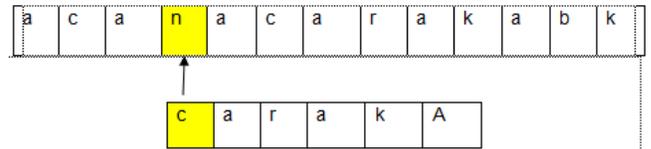


5. Bandingkan *pattern* [3] dan *string* [4]



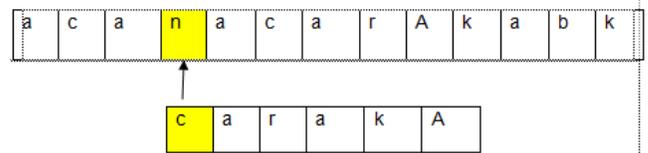
*Pattern* [3] tidak cocok dengan *string* [4]. Karena ada tidak ada kecocokan, maka algoritma *Knuth Morris Pratt* akan menggunakan informasi yang telah disimpan pada langkah dua dan tiga untuk menentukan pergeseran selanjutnya. Sehingga pencocokan yang dilakukan bukanlah antara *pattern* [1] dengan *string* [3] seperti jika menggunakan algoritma *brute force*, namun antara *pattern* [1] dengan *string* [4] sehingga dapat mempercepat proses pencocokan *string*.

6. Bandingkan *pattern* [1] dan *string* [4]



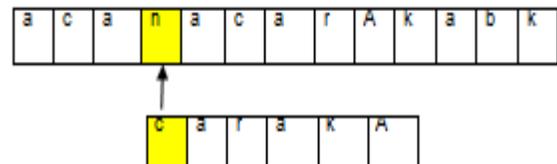
*Pattern* [1] tidak cocok dengan *string* [4]. Maka *pattern* akan bergeser satu posisi ke kanan.

7. Bandingkan *pattern* [1] dan *string* [5]



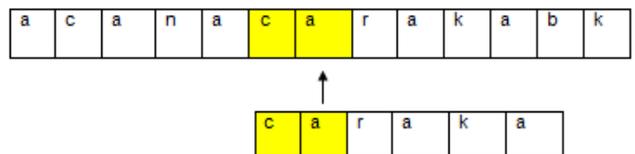
*Pattern* [1] tidak cocok dengan *string* [5]. Maka *pattern* akan bergeser satu posisi ke kanan.

8. Bandingkan *pattern* [1] dan *string* [5]

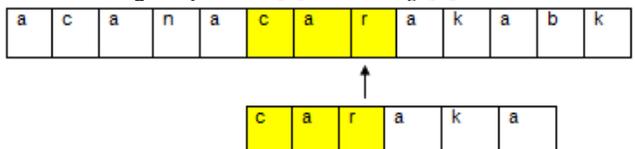


*Pattern* [1] cocok dengan *string* [6]. Karena ada kecocokan, maka algoritma *Knuth Morris Pratt* akan menyimpan informasi ini, dan *pattern* tidak akan melakukan pergeseran dan melanjutkan pencocokan *pattern* [2] dengan *string* [7].

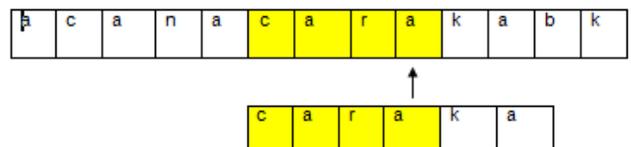
9. Bandingkan *pattern* [2] dan *string* [7]



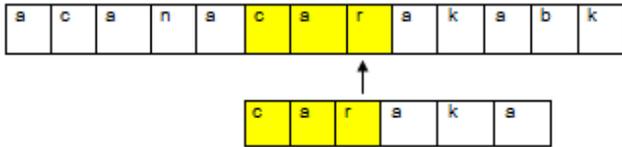
10. Bandingkan *pattern* [3] dan *string* [8]



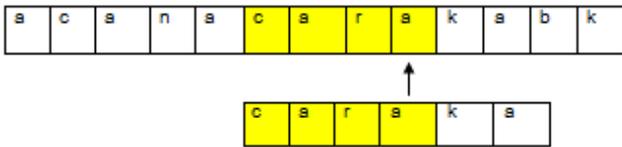
11. Bandingkan *pattern* [4] dan *string* [9]



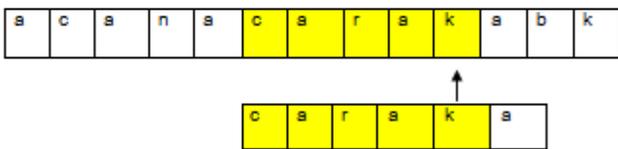
10. Bandingkan *pattern* [3] dan *string* [8]



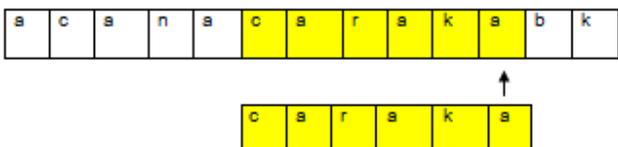
11. Bandingkan *pattern* [4] dan *string* [9]



12. Bandingkan *pattern* [5] dan *string* [10]



13. Bandingkan *pattern* [6] dan *string* [11]



**IV. HASIL DAN PEMBAHASAN**

**4.1 Hasil**

a. Menu Tampilan Index

Menu Tampilan index merupakan tampilan awal dalam aplikasi, adapun Menu tampilan index dapat dilihat pada gambar 4.1.



Gambar 4.1 Rancangan Tampilan Index

b. Menu Tampilan Pos

Menu Tampilan Pos merupakan tampilan informasi pos dalam aplikasi, adapun Menu tampilan pos dapat dilihat pada gambar 4.2.



Gambar 4. 2 Rancangan Tampilan Pos

c. Menu Tampilan Peraturan

Menu Tampilan Pos merupakan tampilan informasi pos dalam aplikasi, adapun rancangan tampilan pos dapat dilihat pada gambar 4. 3.



Gambar 4.3 Menu Tampilan Peraturan

d. Menu Tampilan Buku

Menu Tampilan Buku merupakan tampilan informasi data buku dalam aplikasi, adapun Menu Tampilan Buku dapat dilihat pada gambar 4.4.



Gambar 4.4 Menu Tampilan Buku

e. Menu Tampilan Login

Menu Tampilan Login merupakan tampilan input data admin pada data perpustakaan, adapun Menu Tampilan administrator dapat dilihat pada gambar 4.5.



Gambar 4.5 Rancangan Tampilan Login

f. Menu Tampilan Petugas

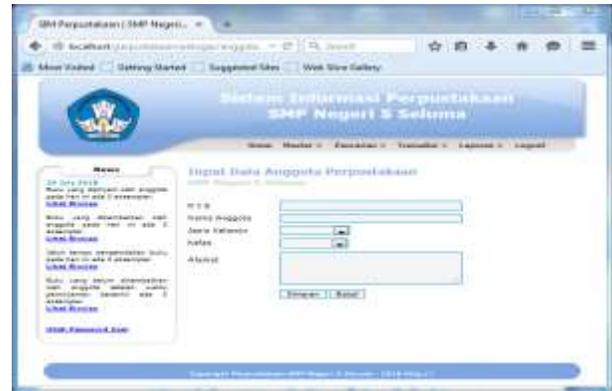
Menu Tampilan Petugas merupakan tampilan informasi data buku dalam aplikasi, adapun Menu Tampilan Petugas dapat dilihat pada gambar 4.6.



Gambar 6 Rancangan Tampilan Petugas

g. Menu Tampilan Master Anggota

Menu Tampilan Master Anggota merupakan tampilan informasi data buku dalam aplikasi, adapun Menu Tampilan Master Anggota dapat dilihat pada gambar 4.7.



Gambar 7 Rancangan Tampilan Master Anggota

h. Menu Tampilan Master Buku

Menu Tampilan Master Buku merupakan tampilan informasi data buku dalam aplikasi, adapun Menu Tampilan Master Buku dapat dilihat pada gambar 4.8.



Gambar 4.8 Rancangan Tampilan Master Buku

i. Menu Tampilan Pencarian Anggota

Menu Tampilan Pencarian Anggota merupakan tampilan informasi data buku dalam aplikasi, adapun Menu Tampilan Pencarian Anggota dilihat pada gambar 4.9.



Gambar 9 Rancangan Tampilan Pencarian Anggota



- [8] Syaroni, Muhammad. 2012. Pencocokan String Berdasarkan Kemiripan Ucapan (Phonetic String Matching) dalam Bahasa Inggris. Institut Teknologi Bandung.
- [9] Kurniawati, A., Wicaksana, I., & Sekarwati, K. A. (2012). Arsitektur Untuk Aplikasi Deteksi Kesamaan Dokumen Bahasa Indonesia. *Konferensi Nasional Sistem Informasi 2012, STMIK-STIKOM Bali 23-25 Pebruari 2012*, (073).
- [10] Astuti, W., 2017. Analisis String Matching Pada Judul Buku dengan Algoritma Knuth-Morris Pratt (KMP). *ILKOM Jurnal Ilmiah*, 9(2), pp.167-172.
- [11] Suparwanta, Agus, and Riyadi J. Iskandar. "PENERAPAN ALGORITMA RABIN KARP UNTUK MEDETEKSI KEMIRIPAN DUA DOKUMEN TEKS." *INTEKSIS* 4, no. 1 (2017).
- [12] Mesran, M. (2017). IMPLEMENTASI ALGORITMA BRUTE FORCE DALAM PENCARIAN DATA KATALOG BUKU PERPUSTAKAAN. *Majalah Ilmiah INTI (Informasi dan Teknologi Ilmiah)*, 3(1).