

# Implementasi *Text Mining* Pada Aplikasi Pengarsipan Berbasis Web Menggunakan Algoritma *Naïve Bayes*

Yungdo Dolot<sup>1</sup>, Parabelem Tinno Dolf Rompas<sup>2</sup>, Vivi Peggie Rantung<sup>3</sup>

<sup>1,2,3</sup> Program Studi Teknik Informatika, Fakultas Teknik, Universitas Negeri Manado

Correspondent Author :

[18210047@unima.ac.id](mailto:18210047@unima.ac.id)

**Abstract** — Archive management in the Technology and Vocational Education Study Program found several obstacles in searching for letter data. With the condition that searching for letter data is less effective and takes quite a long time for lecturers just to search for one letter of data, a website-based archive management system was designed which makes it easier for lecturers to search for data. This website implements text mining using the Naïve Bayes algorithm. The main aim of this research is to make it easier to manage and search for incoming and outgoing mail data in a web-based archiving application. This system was successfully developed using the prototype development method. Tests carried out using the black box testing method obtained good results, where 22 test points were successfully achieved. The Naive Bayes algorithm is used to classify letter data, where out of the 20 tests carried out, all the tests get results that match what was expected. These results show that the document search efficiency and classification accuracy are significantly improved. Therefore, this research shows the success of Text Mining using the Naive Bayes algorithm in web-based archiving applications.

**Keyword** — Naïve Bayes, Text Mining, Archiving, Prototype, Website.

**Abstrak** — Pengelolaan arsip di Program Studi Pendidikan Teknologi dan Kejuruan terdapat beberapa kendala dalam pencarian data surat. Dengan kondisi pencarian data surat yang kurang efektif dan memakan waktu dosen yang cukup lama hanya untuk mencari satu data surat, maka dirancanglah sebuah sistem manajemen arsip berbasis website yang memudahkan dosen dalam pencarian data. Website ini mengimplementasikan *text mining* menggunakan algoritma Naïve Bayes. Tujuan utama dari penelitian ini adalah untuk mempermudah dalam pengelolaan dan pencarian data surat masuk dan surat keluar di dalam aplikasi pengarsipan berbasis web. Sistem ini berhasil dikembangkan menggunakan metode pengembangan prototipe. Pengujian yang dilakukan menggunakan metode pengujian *black box* mendapatkan hasil yang baik, dimana 22 poin pengujian berhasil dicapai. Algoritma Naive Bayes digunakan untuk mengklasifikasikan data surat yang dimana dari 20 pengujian yang dilakukan, dari keseluruhan pengujian mendapatkan hasil yang sesuai dari apa yang diharapkan. Hasil ini menunjukkan bahwa efisiensi pencarian dokumen dan akurasi klasifikasi meningkat secara signifikan. Oleh karena itu penelitian ini

menunjukkan keberhasilan *Text Mining* menggunakan algoritma Naive Bayes dalam aplikasi pengarsipan berbasis web.

**Kata kunci** — Naïve Bayes, Text Mining, Pengarsipan, Prototipe, Website.

## I. PENDAHULUAN

Teknologi saat ini telah berkembang secara pesat dan terus berevolusi hingga sekarang, hal ini dapat dibuktikan dengan banyaknya inovasi dan penemuan yang sederhana hingga rumit. Seperti penemuan teknologi website masa kini yang bermacam-macam.

Kemajuan teknologi yang sangat pesat saat ini tentunya menuntut berbagai instansi agar tetap mengikuti perkembangan. Upaya yang dilakukan guna mengikuti perkembangan yaitu jaringan komputer khususnya internet. Salah satunya adalah penerapan aplikasi pengarsipan itu yang sangat dibutuhkan oleh Program studi Pendidikan Teknologi dan Kejuruan Program Pascasarjana Unima dalam mengelola data kearsipan. *E-learning* merupakan sebuah teknologi untuk memberikan informasi berupa solusi yang bisa meningkatkan pemahaman, pengetahuan, dan kemampuan dalam memberikan informasi dalam bentuk yang jelas.

Kearsipan adalah proses mulai dari penciptaan, penerimaan, pengumpulan, pengaturan, pengendalian, pemeliharaan dan perawatan serta penyimpanan data menurut sistem tertentu, sehingga saat dibutuhkan dapat cepat dan tepat ditemukan. Proses pengarsipan di bidang pendidikan khususnya pada lingkup Program Studi Pendidikan Teknologi dan Kejuruan Pogram Pascasarjana Unima merupakan suatu proses yang mendukung kinerja Prodi dalam bidang pencarian data surat.

## II. LANDASAN TEORI

### A. *Text Mining*

*Text Mining* adalah proses menemukan informasi dalam kumpulan teks besar, dan secara otomatis mengidentifikasi pola dan hubungan yang menarik dalam data tekstual [1]. *Text Mining* memiliki definisi menambang data yang berupa teks dimana sumber data biasanya didapatkan dari dokumen, dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antara dokumen [2].

### B. Pengarsipan

Pengarsipan adalah suatu proses mulai dari penciptaan, penerimaan, pengumpulan, pengaturan, pengendalian, pemeliharaan, dan perawatan serta penyiapan arsip menurut sistem tertentu. Pengarsipan adalah surat pekerjaan kantor atau pekerjaan tata usaha yang banyak di lakukan oleh setiap badan usaha baik dalam pemerintahan maupun usaha swasta [3].

### C. Website

Website adalah kumpulan dari halaman-halaman situs, yang terangkum dalam sebuah domain atau subdomain, yang tempatnya berada di dalam World Wide Website ( WWW ) di dalam internet, Website juga dapat diartikan sebagai sebuah halaman yang berisi data, baik data teks, gambar, suara dan lainnya yang dapat diakses secara *online* [4].

### D. Naïve Bayes

Naïve Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari data set yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas [5].

## III. METODE PENELITIAN

Dalam melakukan pengumpulan data, penulis melakukan tiga metode pengumpulan data:

### a. *Observation* (Pengamatan)

Penulis melakukan pengamatan langsung dilokasi penelitian yaitu di Program Studi Pendidikan Teknologi dan Kejuruan Program Pascasarjana Unima.

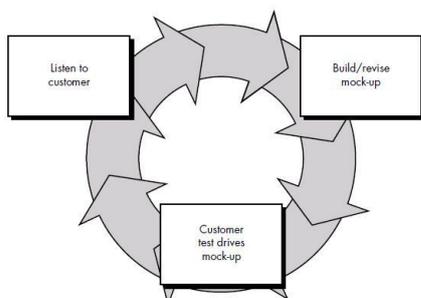
### b. *Interview* (Wawancara)

Penulis mengumpulkan data dan informasi secara lengkap dan jelas, maka penulis melakukan wawancara langsung dengan pimpinan prodi mengenai masalah yang ada berhubungan dengan sistem informasi kearsipan.

### c. *Literature* (Studi pustaka)

Mengumpulkan referensi yang bersumber dari buku dan jurnal yang mendukung penelitian.

Tahap penelitian yang digunakan yaitu dengan menggunakan metode *Prototype* yang ditampilkan pada gambar 1.



Gambar 1. Metode Prototype

Berikut ini adalah tahapan pengembangan sistem metode *prototype* :

### A. Mendengarkan Pelanggan

Pada tahap pertama dari metode ini, penulis melakukan menganalisis kebutuhan. Tahap yang dilakukan antara lain:

- Analisis sistem yang berjalan
- Mengidentifikasi masalah yang terjadi
- Mengusulkan penyelesaian masalah
- Analisis kebutuhan fungsional

### B. Membangun/Memperbaiki Prototipe

Setelah mengetahui sistem yang akan dikembangkan, selanjutnya akan dilakukan perancangan dan pembuatan sistem.

#### a. Perancangan Sistem

Sistem yang akan dikembangkan akan dirancang menggunakan *Unified Modelling Language (UML)*

#### b. Pengkodean

Pada tahap ini, akan dilakukan pembuatan program menggunakan bahasa pemrograman Javascript dengan Framework ExpressJS untuk database yang akan diterapkan kedalam sistem yang akan dikembangkan.

#### c. Pembuatan Prototipe

Pada tahap ini, akan lakukan pembuatan prototipe pada sistem yang akan dikembangkan yaitu dengan membuat *interface* secara langsung berdasarkan perancangan sistem.

### C. Pengujian Prototipe

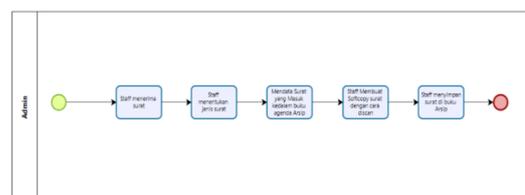
Pada tahap ini, akan dilakukan evaluasi oleh pelanggan untuk mendapatkan tanggapan dari pelanggan untuk mengetahui apakah sistem yang dibuat sudah sesuai dengan apa yang diharapkan atau belum.

## IV. HASIL DAN PEMBAHASAN

### A. Mendengarkan Pelanggan

Analisis sistem yang berjalan

Saat ini proses manajemen arsip yang dilakukan masih secara konvensional/manual mulai dari membuat dan menyimpan arsip data surat.



Gambar 2. Proses Bisnis Manual

Pada Gambar 2 merupakan proses manual manajemen arsip yang ada pada Program Studi Pendidikan Teknologi dan Kejuruan Program Pascasarjana Unima. Proses ini dimulai dengan staff menerima surat kemudian staff menentukan

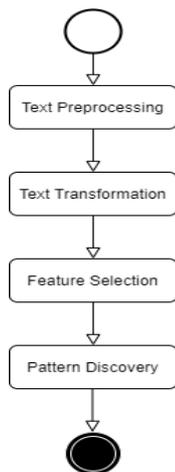
jenis surat apakah surat masuk atau surat keluar selanjutnya staff mendata surat yang masuk kedalam buku agenda arsip kemudian staff membuat *softcopy* surat dengan cara discan setelah itu staff menyimpan surat dibuku arsip.

Mengidentifikasi masalah yang terjadi

- Proses manajemen arsip masih melakukan cara manual
- Staff harus menentukan jenis surat yang masuk

Mengusulkan penyelesaian masalah

Dengan melihat permasalahan yang ada, dibutuhkan sistem informasi berbasis Website yang dapat membantu dalam proses manajemen Arsip di Program Studi Pendidikan Teknologi dan Kejuruan Program Pascasarjana Unima. Dalam pengembangannya peneliti menggunakan *text mining* dengan algoritma naïve bayes sebagai algoritma yang akan digunakan dalam sistem. Dalam implementasi algoritma naïve bayes peneliti menggunakan *library* WhichX dimana *library* ini umumnya digunakan untuk implementasi *text mining* pada aplikasi yang menggunakan nodeJS sebagai *runtime*.



Gambar 3. Alur *Text Mining* Sistem

Gambar 3. Menunjukkan alur *text mining system*, pertama tahap *text preprocessing* untuk melakukan analisis semantik dan sintaktik terhadap teks. Selanjutnya tahap *text transformation* yaitu pembentukan atribut mengacu pada proses untuk mendapatkan gambaran dokumen yang diharapkan. Selanjutnya tahap *feature selection* dalam sistem ini tahap ini bertujuan untuk mengurangi dimensi pada tahap *text transformation* menggunakan operasi *stop words removal* dan *stemming*. Selanjutnya yang terakhir tahap *pattern discovery* merupakan tahap untuk menemukan pola dari keseluruhan text, tahap ini disebut juga dengan *text mining*. Dalam sistem ini atribut yang digunakan untuk dilakukan *text mining* yaitu perihal surat.



Gambar 4. Flowchart *Text Mining* Sistem

Gambar 4. Menunjukkan alur *text mining system*, pertama admin melakukan *upload* surat melalui halaman surat, jika sistem gagal melakukan validasi surat maka sistem akan mengarahkan admin ke halaman surat. Namun jika berhasil sistem akan melakukan tahapan *text mining* seperti yang ditunjukkan pada gambar 3. Terhadap surat yang di-*upload* admin berdasarkan perihal surat. Setelah melakukan *text mining system* melakukan klasifikasi surat menggunakan algoritma Naïve Bayes. Setelah melakukan klasifikasi sistem akan menghasilkan *output* apakah surat yang di-*upload* termasuk surat masuk atau surat keluar.

Dalam proses klasifikasi terdapat dua kategori surat yang merupakan *output* dari proses *text mining* yaitu surat masuk dan surat keluar dengan kata kunci dari masing-masing kategori surat sebagai berikut:

Tabel 1. Kata Kunci Surat

Kategori	Kata kunci
Surat masuk	Pemberitahuan, permohonan, monitoring, undangan, dan edaran.
Surat keluar	Permohonan ujian, permohonan aktif kuliah, usulan, persetujuan, penerbitan, dan pengantar.

Dikarenakan pada kedua kategori surat terdapat kata kunci yang sama yaitu permohonan maka kata kunci permohonan diberikan pada kedua kategori surat dengan memberi pembeda berupa spasi sebagai penghubung kata untuk kategori surat keluar. Sehingga sistem dapat dengan akurat menentukan kategori antara surat masuk dan surat keluar untuk kata kunci permohonan.

Analisis kebutuhan fungsional

Kebutuhan fungsional yaitu mendeskripsikan fitur dan fungsi yang disediakan atau diberikan oleh sistem bagi pengguna. Berikut merupakan hasil analisis kebutuhan fungsional terhadap sistem yang akan dikembangkan :

- Halaman login untuk admin
- Upload data surat
- Kelola data surat

d. Mencari data surat

B. Membangun/Memperbaiki Prototipe

Berikut merupakan rancangan dari sistem yang akan dikembangkan. Pengembangan sistem akan menggunakan pemodelan UML.

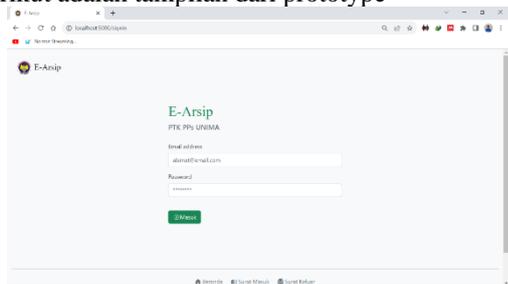
- 1. Perancangan Sistem
  - a. Penentuan Aktor
  - b. Definisi Usecase
  - c. Usecase Skenario
  - d. Usecase Diagram
  - e. Activity Diagram
  - f. Sequence Diagram

2. Pengkodean

Pengembang menggunakan NodeJS sebagai server lokal, MongoDB untuk database, Framework ExpressJS, Visual Studio Code untuk teks editor dan Google Chrome untuk browser.

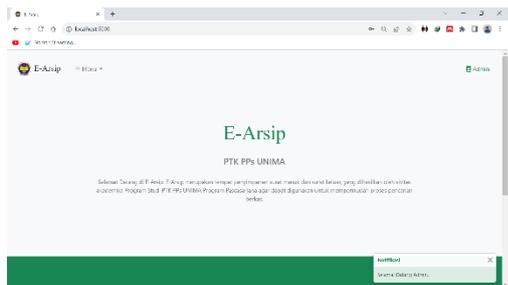
3. Pembuatan Prototipe

Berikut adalah tampilan dari prototipe



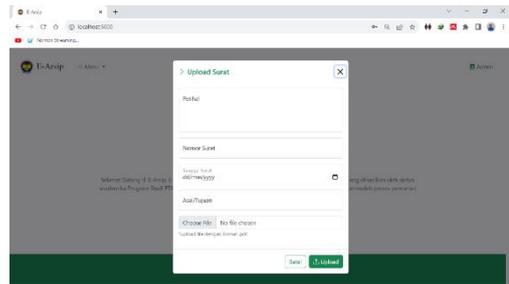
Gambar 5. Halaman Login

Gambar 5. Adalah tampilan halaman login admin.



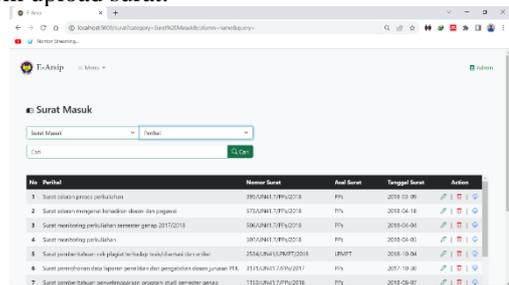
Gambar 6. Halaman Utama

Gambar 6. Adalah tampilan halaman utama dimana menampilkan bagian utama dari sistem.



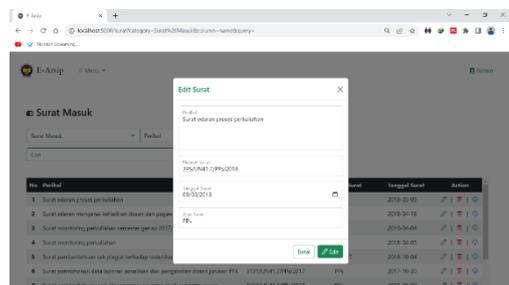
Gambar 7. Upload Surat

Gambar 7. Upload surat dimana menampilkan detail form upload surat.



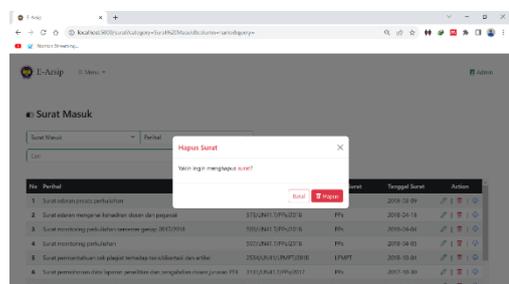
Gambar 8. Cari Data Surat Masuk

Gambar 8. Adalah halaman cari data surat dimana menampilkan daftar data surat masuk.



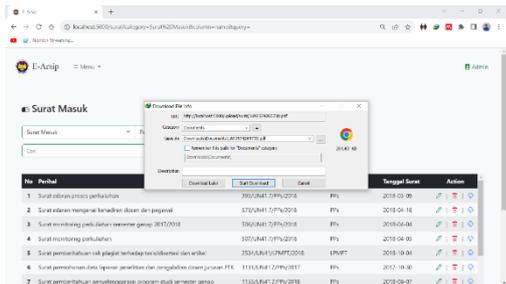
Gambar 9. Edit Surat Masuk

Gambar 9. Edit surat dimana menampilkan detail dari form edit surat masuk.

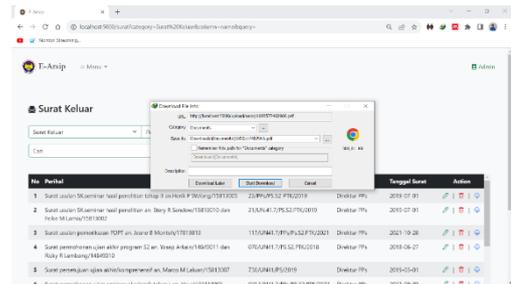


Gambar 10. Hapus Surat Masuk

Gambar 10. Hapus surat dimana menampilkan form dari hapus data surat masuk.



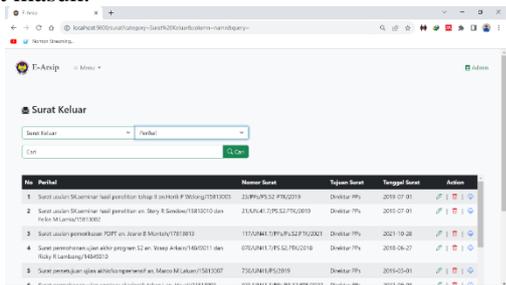
Gambar 11. Download Surat Masuk



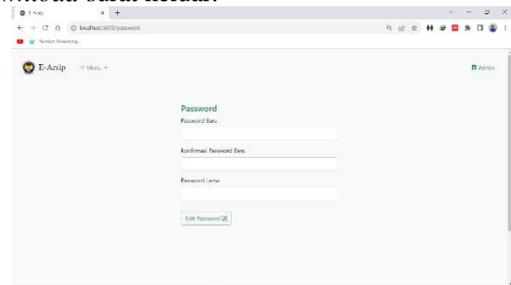
Gambar 15. Download Surat Keluar

Gambar 11. Menampilkan tampilan untuk men-download surat masuk.

Gambar 15. Menampilkan tampilan untuk men-download surat keluar.



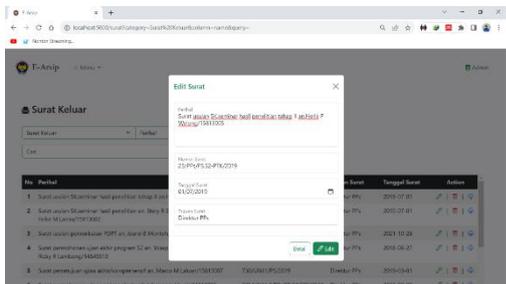
Gambar 12. Cari Data Surat Keluar



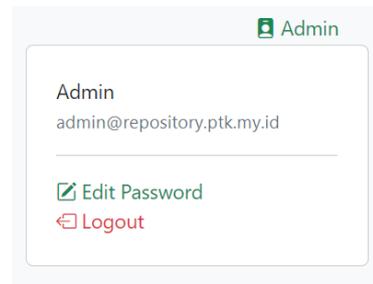
Gambar 16. Edit Password

Gambar 12. Adalah halaman cari data surat dimana menampilkan daftar data surat Keluar.

Gambar 16. Edit password dimana menampilkan detail halaman edit password.



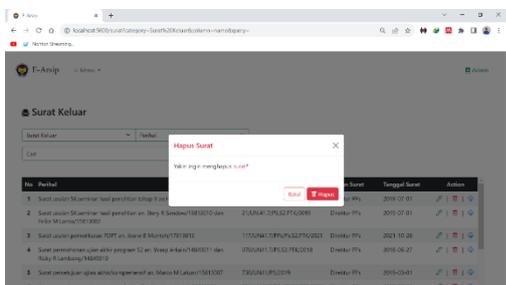
Gambar 13. Edit Surat Keluar



Gambar 17. Logout

Gambar 13. Edit surat dimana menampilkan detail dari form edit surat Keluar.

Gambar 17. Menampilkan form untuk logout



Gambar 14. Hapus Surat Keluar

Gambar 14. Hapus surat dimana menampilkan form dari hapus data surat keluar.

### C. Pengujian Prototipe

Berikut merupakan pengujian prototipe menggunakan metode *Black Box Testing*.

Tabel 2. Black Box Testing

No	Tujuan yang ingin dicapai	Input	Output yang di harapkan	Has il
1	Menampilkan Halaman Login	Mengunjungi Websites E-Arsip dengan memasuk	Sistem menampilkan halaman login	Sukses

		kan <i>link Website</i>		
2	Melakukan Login	Memasukkan <i>username</i> dan <i>password</i> yang sesuai	Sistem menampilkan halaman utama	Sukses
3	Menampilkan Halaman Utama	Melakukan <i>login</i> ke dalam sistem menggunakan akun yang sesuai	Sistem menampilkan halaman utama	Sukses
4	Menampilkan <i>Form Upload Surat</i>	Menekan tombol <i>Upload Surat</i> pada menu yang terdapat pada navigasi	Sistem menampilkan form <i>upload surat</i>	Sukses
5	Meng- <i>upload</i> Surat	Memasukkan data surat yang sesuai dengan <i>form</i> yang ditampilkan lalu menekan tombol <i>Upload</i>	Sistem menyimpan data surat sesuai dengan yang di- <i>input</i> -kan Admin ke dalam form	Sukses
6	Menampilkan Halaman Surat Masuk	Menekan tombol Surat Masuk pada menu yang terdapat pada navigasi	Sistem menampilkan halaman surat masuk	Sukses
7	Mencari Surat Masuk	Mengisi <i>form</i> pencarian pada halaman Surat	Sistem menampilkan Surat Masuk sesuai dengan data	Sukses

		Masuk dengan sesuai dan menekan tombol Cari	yang diinput oleh Admin	
8	Menampilkan Form <i>Edit Surat Masuk</i>	Menekan tombol <i>Edit Surat Masuk</i> pada tabel yang terdapat pada halaman Surat Masuk	Sistem menampilkan form <i>edit surat masuk</i>	Sukses
9	<i>Edit Surat Masuk</i>	Memasukkan data surat yang sesuai dengan <i>form</i> yang ditampilkan lalu menekan tombol <i>Edit</i>	Sistem menyimpan data surat yang baru sesuai dengan yang di- <i>input</i> -kan Admin ke dalam form	Sukses
10	Menampilkan Form Hapus Surat Masuk	Menekan tombol Hapus Surat Masuk pada tabel yang terdapat pada halaman Surat Masuk	Sistem menampilkan form hapus surat masuk	Sukses
11	Hapus Surat Masuk	Menekan tombol hapus yang terdapat pada <i>form</i> hapus Surat Masuk	Sistem Menghapus data surat sesuai dengan yang dipilih Admin dari database	Sukses

12	<i>Download</i> Surat Masuk	Menekan tombol <i>download</i> yang terdapat pada halaman surat masuk	Admin mengunduh surat masuk sesuai dengan yang dipilih Admin	Sukses
13	Menampilkan Halaman Surat Keluar	Menekan tombol Surat Keluar pada menu yang terdapat pada navigasi	Sistem menampilkan halaman surat Keluar	Sukses
14	Mencari Surat Keluar	Mengisi <i>form</i> pencarian pada halaman Surat Keluar dengan sesuai dan menekan tombol Cari	Sistem menampilkan Surat Keluar sesuai dengan data yang diinput oleh Admin	Sukses
15	Menampilkan <i>Form</i> Edit Surat Keluar	Menekan tombol Edit Surat Keluar pada tabel yang terdapat pada halaman Surat Keluar	Sistem menampilkan <i>form edit</i> surat keluar	Sukses
16	<i>Edit</i> Surat Keluar	Memasukkan data surat yang sesuai dengan form yang ditampilkan lalu menekan	Sistem menyimpan data surat yang baru sesuai dengan yang di-input-kan Admin ke dalam form	Sukses

		tombol <i>Edit</i>		
17	Menampilkan <i>Form</i> Hapus Surat Keluar	Menekan tombol Hapus Surat Keluar pada tabel yang terdapat pada halaman Surat Keluar	Sistem menampilkan <i>form</i> hapus surat keluar	Sukses
18	Hapus Surat Keluar	Menekan tombol hapus yang terdapat pada <i>form</i> hapus Surat Keluar	Sistem Menghapus data surat sesuai dengan yang dipilih Admin dari database	Sukses
19	<i>Download</i> Surat Keluar	Menekan tombol <i>download</i> yang terdapat pada halaman surat keluar	Admin mengunduh surat keluar sesuai dengan yang dipilih Admin	Sukses
20	Menampilkan Halaman <i>Edit Password</i>	Menekan tombol <i>Edit Password</i> pada menu yang terdapat pada navigasi	Sistem menampilkan halaman <i>edit password</i>	Sukses
21	<i>Edit Password</i>	Memasukkan data sesuai dengan halaman <i>Edit Password</i> yang ditampilkan lalu menekan tombol <i>Edit</i>	Sistem menyimpan data yang baru sesuai dengan yang di-input-kan Admin ke dalam halaman <i>edit password</i>	Sukses

		<i>Password</i>		
22	<i>Logout</i>	Menekan tombol <i>Logout</i> pada menu yang terdapat pada navigasi	Sistem menampilkan halaman login	Sukses

14	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai
15	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai
16	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai
17	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai
18	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai
19	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai
20	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai

### Pengujian Algoritma Naïve Bayes

Dalam pengujian ini, dilakukan pengujian dengan input 10 Surat Masuk dan 10 Surat Keluar, dengan total pengujian 20 pengujian.

Tabel 3. Pengujian Algoritma Naïve Bayes

Pengujian Ke-	Input	Output	Hasil yang diharapkan	Sesuai/Tidak Sesuai
1	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
2	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
3	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
4	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
5	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
6	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
7	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
8	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
9	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
10	Surat Masuk	Surat Masuk	Surat Masuk	Sesuai
11	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai
12	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai
13	Surat Keluar	Surat Keluar	Surat Keluar	Sesuai

### VII. KESIMPULAN

Dalam pengembangan sistem E-Arsip di program studi Pendidikan Teknologi dan Kejuruan Program Pascasarjana Unima berbasis website menggunakan metode *prototype*, dapat diambil kesimpulan sebagai berikut:

1. Sistem arsip di program studi Pendidikan Teknologi dan Kejuruan Program Pascasarjana Unima berbasis website berhasil dikembangkan menggunakan metode pengembangan prototipe.
2. Pengujian yang dilakukan menggunakan metode pengujian black box mendapatkan hasil yang baik, dimana 22 poin pengujian berhasil dicapai.

### DAFTAR ACUAN

- [1] A. F. Firdaus and W. I. Firdaus, "Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi:(Sebuah Ulasan)," *JUPITER (Jurnal Penelit. Ilmu dan Tek. Komputer)*, vol. 13, no. 1, pp. 66–78, 2021.
- [2] A. Nurzahputra and M. A. Muslim, "Analisis sentimen pada opini mahasiswa menggunakan natural language processing," in *Seminar Nasional Ilmu Komputer (SNIK 2016)*, 2016, pp. 114–118.
- [3] dan A. Donni, Gamida, "Sistem Penyimpanan Arsip," *Angew. Chemie Int. Ed. 6(11)*, 951–952., pp. 8–24, 2013.
- [4] A. Josi, "Penerapan metode prototyping dalam pembangunan website desa (studi kasus desa sugihan kecamatan rambang)," *J. Teknol. Inf. Mura*, vol. 9, no. 1, 2017.
- [5] R. Y. Hayuningtyas, "Penerapan Algoritma Naïve Bayes untuk Rekomendasi Pakaian Wanita," *J. Inform.*, vol. 6, no. 1, pp. 18–22, 2019.