

Pengembangan Sistem Informasi Pengarsipan Surat Masuk dan Surat Keluar Berbasis Web dengan Fitur Klasifikasi Otomatis Menggunakan Algoritma Naive Bayes di SMP Negeri 5 Sukoharjo

Sri Rejeki^{1*}, Sri Sumarlinda², Eko Purwanto³

¹ Program Studi Sistem Informasi,

² Program Studi Sistem Informasi,

³ Program Studi Sistem Informasi,

Universitas Duta Bangsa Surakarta

Universitas Duta Bangsa Surakarta

Universitas Duta Bangsa Surakarta

^{1*}220101146@mhs.udb.ac.id

²sri_sumarlinda@udb.ac.id

³eko_purwanto@udb.ac.id

Abstrak---Pengelolaan arsip surat merupakan salah satu kegiatan administrasi yang memiliki peran penting dalam menjaga ketertiban dokumen pada suatu instansi pendidikan. Di SMP Negeri 5 Sukoharjo, proses pengarsipan surat masuk dan surat keluar masih dilakukan secara manual sehingga menimbulkan berbagai permasalahan, seperti kesulitan dalam pencarian dokumen, tingginya risiko kehilangan arsip, serta proses pengelompokan surat yang membutuhkan waktu cukup lama. Penelitian ini bertujuan mengembangkan sistem informasi pengarsipan surat berbasis web yang dilengkapi dengan fitur klasifikasi otomatis menggunakan algoritma Naïve Bayes untuk membantu proses pengelompokan surat secara lebih cepat dan akurat. Metode pengembangan sistem yang digunakan adalah Waterfall yang terdiri atas tahapan analisis kebutuhan, perancangan sistem, implementasi, pengujian, dan pemeliharaan. Sistem dibangun menggunakan bahasa pemrograman Python dengan framework Flask, basis data SQLite, serta Bootstrap sebagai antarmuka pengguna. Algoritma Naïve Bayes diterapkan untuk mengklasifikasikan kategori surat berdasarkan isi perihal surat sehingga membantu proses pengelompokan arsip secara otomatis. Hasil penelitian menunjukkan bahwa sistem mampu mempermudah proses pengelolaan surat mulai dari penyimpanan, pencarian, pengubahan, penghapusan hingga pencetakan laporan dalam bentuk PDF dan Excel. Selain itu, penerapan algoritma Naïve Bayes dapat membantu proses klasifikasi surat menjadi lebih cepat, efektif, dan terstruktur sehingga meningkatkan efisiensi administrasi pengarsipan surat di SMP Negeri 5 Sukoharjo.

Kata Kunci----Sistem Informasi, Pengarsipan Surat, Naïve Bayes, Flask, SQLite.

Abstract---Mail archive management is an essential administrative activity that supports document organization in educational institutions. At SMP Negeri 5 Sukoharjo, incoming and outgoing mail archiving is still carried out manually, resulting in difficulties in document retrieval, the risk of document loss, and time-consuming classification processes. This study aims to develop a web-based mail archiving information system equipped with automatic classification features using the Naïve Bayes algorithm. The system was developed using the Waterfall software development method, which includes requirement analysis, system design, implementation, testing, and maintenance. The system was developed using Python with the Flask framework, SQLite database, and Bootstrap for the user interface. The Naïve Bayes algorithm was implemented to classify mail categories automatically based on the subject of each letter. The results indicate that the developed system facilitates mail management activities, including storing, searching, editing, deleting, and generating reports in PDF and Excel formats. Furthermore, the implementation of the Naïve Bayes algorithm improves the efficiency and effectiveness of document classification, resulting in a more organized and efficient mail archiving process at SMP Negeri 5 Sukoharjo.

Keywords----Information System, Mail Archive, Naive Bayes, Flask, SQLite.

I. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong berbagai instansi untuk memanfaatkan sistem informasi guna meningkatkan efektivitas dan efisiensi proses administrasi. Salah satu aspek administrasi yang memerlukan pengelolaan secara sistematis adalah pengarsipan surat, karena surat merupakan dokumen resmi yang menjadi bukti pelaksanaan kegiatan organisasi. Oleh sebab itu, pengelolaan arsip surat perlu dilakukan secara terstruktur agar dokumen dapat disimpan,

ditemukan kembali, dan dimanfaatkan dengan mudah saat diperlukan[1].

SMP Negeri 5 Sukoharjo merupakan salah satu instansi pendidikan yang setiap hari melakukan pengelolaan surat masuk dan surat keluar sebagai bagian dari kegiatan administrasi sekolah. Berdasarkan hasil observasi dan wawancara, proses pengarsipan surat masih dilakukan secara manual melalui menyimpan dokumen fisik maupun file sederhana. Kondisi tersebut menyebabkan proses pencarian arsip memerlukan waktu yang cukup lama, meningkatkan risiko kehilangan dokumen,

serta menyulitkan proses pengelompokan surat berdasarkan kategorinya [2].

Penerapan sistem informasi berbasis web menjadi solusi untuk mengatasi berbagai permasalahan tersebut. Sistem ini memungkinkan proses penyimpanan, pencarian, pembaruan data, serta penyusunan laporan dilakukan secara terintegrasi. Selain itu, penerapan algoritma klasifikasi pada sistem dapat membantu pengguna mengelompokkan surat secara otomatis sesuai kategori yang telah ditentukan sehingga proses pengarsipan menjadi lebih efektif dan efisien [3].

Salah satu algoritma yang banyak digunakan dalam klasifikasi data teks adalah **Naïve Bayes**. Algoritma ini memiliki keunggulan berupa proses perhitungan yang sederhana, kebutuhan komputasi yang relatif rendah, serta mampu menghasilkan tingkat akurasi yang baik pada proses klasifikasi dokumen teks [4]. Pada penelitian ini, algoritma **Naïve Bayes** digunakan untuk mengklasifikasikan kategori surat berdasarkan informasi pada bagian perihal surat sehingga proses pengelompokan arsip dapat dilakukan secara otomatis.

Meskipun telah banyak penelitian mengenai klasifikasi dokumen menggunakan algoritma **Naïve Bayes**, implementasi algoritma tersebut pada sistem pengarsipan surat berbasis web di lingkungan sekolah masih relatif terbatas. Oleh karena itu, penelitian ini mengidentifikasi algoritma **Naïve Bayes** ke dalam sistem pengarsipan surat untuk mendukung proses klasifikasi kategori surat secara otomatis.

Penelitian ini bertujuan mengembangkan Sistem Informasi Pengarsipan Surat Masuk dan Surat Keluar berbasis web menggunakan **framework Flask** dan **basis data SQLite** dengan menerapkan algoritma **Naïve Bayes** sebagai metode klasifikasi otomatis kategori surat. Sistem yang dikembangkan diharapkan mampu membantu pegawai Tata Usaha dalam mengelola arsip surat secara lebih cepat, akurat, dan terstruktur sehingga dapat meningkatkan kualitas pelayanan administrasi di SMP Negeri 5 Sukoharjo [5]-[8].

II. METODOLOGI PENELITIAN

Penelitian ini dilaksanakan di **SMP Negeri 5 Sukoharjo** dengan tujuan mengembangkan Sistem Informasi Pengarsipan Surat Masuk dan Surat Keluar berbasis web yang dilengkapi fitur klasifikasi otomatis menggunakan algoritma **Naïve Bayes**. Metode penelitian yang digunakan meliputi observasi, wawancara, studi pustaka, pengembangan sistem, serta pengujian sistem [9].

Pengumpulan data dilakukan melalui observasi secara langsung terhadap proses pengelolaan surat pada bagian Tata Usaha, wawancara dengan pihak pegawai yang bertanggung jawab terhadap administrasi surat, serta studi pustaka yang bersumber dari buku, jurnal ilmiah, prosiding dan penelitian terdahulu yang berkaitan dengan sistem informasi, pengarsipan surat, text mining, dan algoritma **Naïve Bayes** [10].

Pengembangan sistem menggunakan model Waterfall karena memiliki tahapan yang tersusun secara sistematis sehingga sesuai untuk pembangunan perangkat lunak [11]. Tahapan Waterfall meliputi analisis kebutuhan, perancangan sistem, implementasi, pengujian, dan pemeliharaan. Tahap perancangan menghasilkan desain sistem menggunakan Unified Modeling Language (UML), perancangan basis data, serta rancangan antarmuka pengguna. Implementasi sistem dilakukan menggunakan bahasa pemrograman Python dengan framework Flask, basis data SQLite sebagai media penyimpanan data, serta HTML, CSS, Bootstrap, dan Chart.js sebagai pendukung antarmuka pengguna [6] – [8].

Algoritma **Naïve Bayes** diterapkan untuk mengklasifikasikan kategori surat berdasarkan informasi pada bagian perihal surat. Sebelum proses klasifikasi dilakukan, data teks diproses melalui tahapan **preprocessing** yang meliputi **case folding**, **tokenizing**, **stopword removal**, dan **stemming**. Selanjutnya sistem menghitung probabilitas setiap kategori menggunakan algoritma **Naïve Bayes** sehingga kategori dengan nilai probabilitas tertinggi dipilih sebagai hasil klasifikasi surat [12].

Untuk menghindari nilai probabilitas nol pada proses klasifikasi, perhitungan likelihood dilakukan

menggunakan teknik **Laplace Smoothing** untuk menghindari nilai probabilitas nol pada proses klasifikasi, perhitungan likelihood dilakukan menggunakan teknik Laplace Smoothing sehingga setiap kata tetap memiliki nilai probabilitas. Teknik ini membantu meningkatkan kemampuan model dalam mengklasifikasikan dokumen yang mengandung kata-kata yang belum pernah muncul pada data latih.

Pengujian sistem dilakukan menggunakan metode **Black Box Testing** untuk memastikan setiap fungsi berjalan sesuai kebutuhan pengguna. Pengujian mencakup proses autentikasi pengguna, pengelolaan surat masuk dan surat keluar, pencarian data, unggah dokumen, klasifikasi otomatis menggunakan algoritma **Naïve Bayes**, serta pembuatan laporan dalam format PDF dan Microsoft Excel. Selain itu, performa algoritma dievaluasi menggunakan metode **5-Fold Cross Validation** dan dianalisis melalui **Confusion Matrix** untuk mengetahui tingkat akurasi klasifikasi pada setiap kategori surat[13].

III. HASIL DAN PEMBAHASAN

A. Implementasi Sistem

Hasil penelitian berupa Sistem Informasi Pengarsipan Surat Masuk dan Surat Keluar berbasis web yang dikembangkan menggunakan bahasa pemrograman Python dengan framework **Flask** dan **basis data SQLite**. Sistem dirancang untuk membantu proses administrasi surat di SMP Negeri 5 Sukoharjo mulai dari penyimpanan data, pencarian arsip, klasifikasi kategori surat, hingga pembuatan laporan.

Sistem menerapkan mekanisme autentikasi pengguna melalui halaman login sehingga hanya pengguna yang memiliki hak akses yang dapat mengelola data surat. Setelah berhasil masuk ke sistem, pengguna diarahkan ke halaman dashboard yang menampilkan informasi jumlah surat masuk, surat keluar, kategori surat, serta grafik statistik sebagai ringkasan data arsip.

Fitur utama sistem meliputi pengelolaan data surat masuk dan surat keluar. Pengguna dapat menambahkan, mengubah, menghapus, melihat detail, serta melakukan pencarian data surat berdasarkan nomor surat, tanggal, maupun perihal. Selain itu, sistem mendukung unggah dokumen surat dalam format digital sehingga arsip dapat disimpan secara terpusat dan dapat diakses kembali.

B. Implementasi Algoritma Naïve Bayes

Algoritma **Naïve Bayes** diterapkan untuk melakukan klasifikasi kategori surat secara otomatis berdasarkan data perihal surat yang diinputkan pengguna. Sebelum proses klasifikasi dilakukan, data teks terlebih dahulu melalui tahapan **preprocessing** yang meliputi **case folding, tokenizing, stopword removal, dan stemming**[14]. Tahapan tersebut bertujuan untuk menghasilkan data teks yang lebih terstruktur sehingga meningkatkan kualitas proses klasifikasi.

Selanjutnya sistem menghitung probabilitas setiap kategori menggunakan algoritma **Naïve Bayes** berdasarkan data latih yang telah tersedia. Kategori dengan nilai probabilitas tertinggi dipilih sebagai hasil klasifikasi dan secara otomatis disimpan ke dalam basis data. Dengan adanya proses tersebut, pengguna tidak perlu lagi menentukan kategori surat secara manual sehingga proses pengelompokan arsip menjadi lebih cepat dan konsisten.

Algoritma **Naïve Bayes** diimplementasikan secara manual menggunakan bahasa pemrograman Python tanpa memanfaatkan library machine learning. Model dibangun melalui proses pembentukan vocabulary, perhitungan prior probability berdasarkan jumlah dokumen pada setiap kategori, serta likelihood setiap kata menggunakan **Laplace Smoothing**[15]. Untuk menghindari underflow pada proses perkalian probabilitas, perhitungan posterior dilakukan menggunakan **log probability**[16]. Kategori dengan nilai posterior terbesar dipilih sebagai hasil klasifikasi surat.

C. Hasil Pengujian Sistem

Pengujian sistem dilakukan menggunakan metode **Black Box Testing**[15] untuk memastikan setiap

fungsi berjalan sesuai kebutuhan pengguna. Pengujian dilakukan pada seluruh fitur utama, antara lain proses login, pengelolaan surat masuk, pengelolaan surat keluar, pencarian data, unggah dokumen, klasifikasi otomatis menggunakan algoritma Naïve Bayes, serta pembuatan laporan dalam format PDF dan Excel.

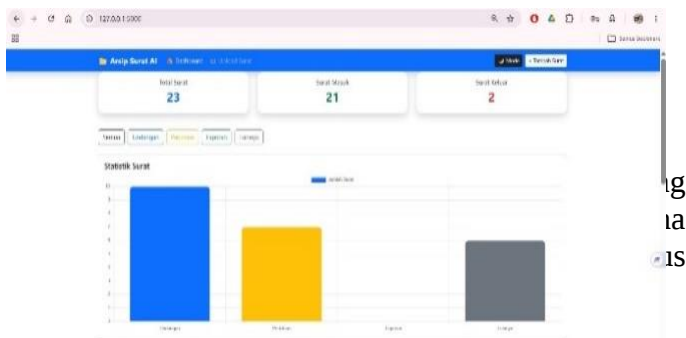
Hasil pengujian menunjukkan bahwa seluruh fungsi sistem dapat dijalankan sesuai dengan skenario pengujian yang telah ditetapkan. Sistem mampu menyimpan data surat, menampilkan informasi secara benar, melakukan pencarian arsip dengan cepat, menghasilkan klasifikasi kategori surat secara otomatis, serta membuat laporan sesuai kebutuhan pengguna.

D. Pembahasan

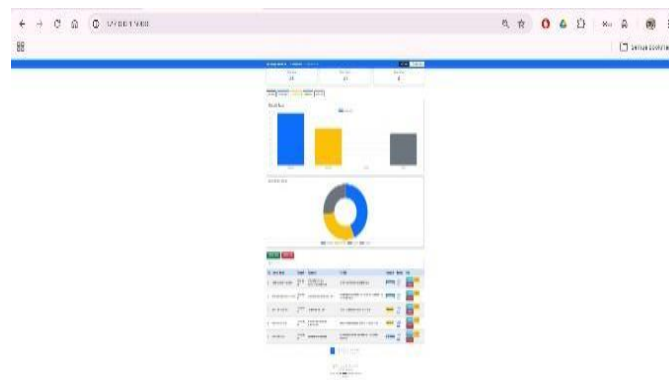
Implementasi sistem informasi pengarsipan surat berbasis web memberikan peningkatan terhadap proses administrasi di SMP Negeri 5 Sukoharjo. Digitalisasi arsip memudahkan proses penyimpanan dan pencarian dokumen dibandingkan metode manual yang sebelumnya digunakan. Selain itu, penerapan algoritma Naïve Bayes mampu membantu proses klasifikasi surat secara otomatis berdasarkan perihal surat sehingga mengurangi kesalahan pengelompokan dan mempercepat proses administrasi.

Fitur pencarian data, pengelolaan arsip digital, visualisasi statistik pada dashboard, serta fasilitas ekspor laporan dalam format PDF dan Excel memberikan kemudahan bagi pengguna dalam melakukan monitoring maupun penyusunan laporan administrasi, dan kerapian pengelolaan arsip surat di lingkungan SMP Negeri 5 Sukoharjo.

A. Gambar hasil implementasi sistem:

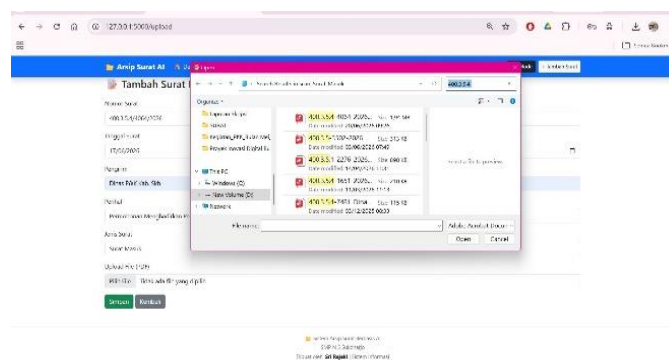


memasukkan username dan password yang valid sehingga keamanan data arsip dapat terjaga.



Gambar 2. Dashboard Sistem

Gambar 2 menunjukkan dashboard sistem yang menyajikan informasi jumlah surat masuk, surat keluar, serta statistic kategori surat. Informasi tersebut membantu pengguna memantau kondisi arsip secara cepat.



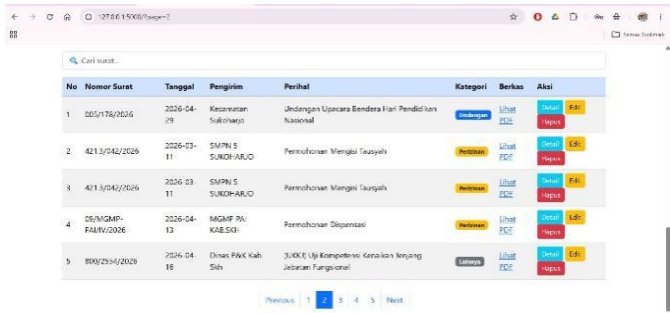
Gambar 3. Halaman Unggah Surat

Gambar 3 menunjuk halaman unggah surat. Pengguna mengisi data surat kemudian mengunggah dokumen dalam format PDF. Setelah data dikirim, sistem secara otomatis melakukan klasifikasi kategori surat menggunakan algoritma Naïve Bayes.



Gambar 4. Hasil Klasifikasi Naïve Bayes

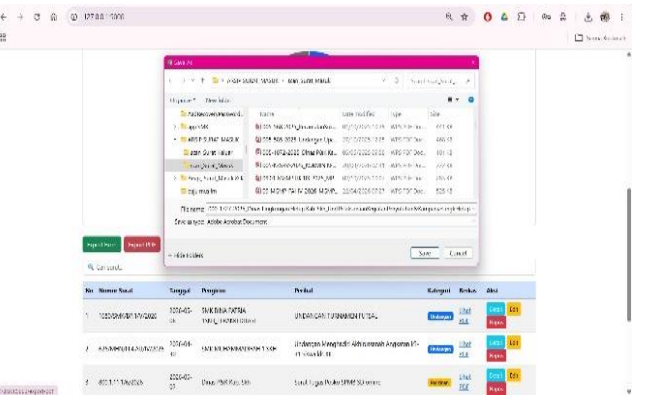
Gambar 4 menunjukkan hasil klasifikasi otomatis menggunakan algoritma Naïve Bayes. Sistem menentukan kategori surat berdasarkan isi perihal yang dimasukkan pengguna sehingga proses pengelompokan arsip menjadi lebih cepat dan konsisten.



No	Nomor Surat	Tanggal	Pengirim	Perihal	Kategori	Berkas	Aksi
1	005/176/2026	2026-04-29	Kecamatan Subaharjo	Undangan Upacara Bendera Hari Pendidikan Nasional	Undangan	Undut PDF	Detail Edit Hapus
2	4213/542/2026	2026-03-11	SMIPA 5 SURABAJA	Pemohonan Mengisi Tawar	Pemohonan	Undut PDF	Detail Edit Hapus
3	4213/542/2026	2026-03-11	SMIPA 5 SURABAJA	Pemohonan Mengisi Tawar	Pemohonan	Undut PDF	Detail Edit Hapus
4	06/MGMP-FARW/2026	2026-04-13	MAGMP PK KASEDI	Pemohonan Dispensasi	Pemohonan	Undut PDF	Detail Edit Hapus
5	100/294/2026	2026-04-16	Dinas PPK Kab Sido	Surat Keterangan Koneksi Jaring Jaringan Fungsional	Surat Keterangan	Undut PDF	Detail Edit Hapus

Gambar 5. Daftar Arsip Surat

Gambar 5 menunjukkan daftar arsip surat yang telah tersimpan pada basis data. Pengguna dapat melihat informasi surat, melakukan pencarian, mengubah data, menghapus data, maupun membuka dokumen yang telah diunggah.

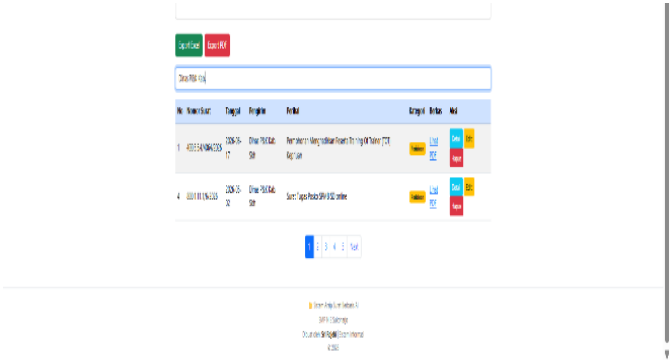


No	Nomor Surat	Tanggal	Pengirim	Perihal	Kategori	Berkas	Aksi
1	100/294/2026	2026-04-16	Dinas PPK Kab Sido	Surat Keterangan Koneksi Jaring Jaringan Fungsional	Surat Keterangan	Undut PDF	Detail Edit Hapus
2	4213/542/2026	2026-03-11	SMIPA 5 SURABAJA	Pemohonan Mengisi Tawar	Pemohonan	Undut PDF	Detail Edit Hapus
3	4213/542/2026	2026-03-11	SMIPA 5 SURABAJA	Pemohonan Mengisi Tawar	Pemohonan	Undut PDF	Detail Edit Hapus

Gambar 6. Ekspor Laporan

Gambar 6 menunjukkan hasil ekspor data arsip dalam format PDF dan Microsoft Excel. Fitur ini memudahkan pengguna dalam menghasilkan laporan administrasi secara cepat, lengkap. Pada pengembangan selanjutnya, sistem masih dapat

dikembangkan agar mendukung akses melalui perangkat bergerak serta notifikasi surat secara otomatis.



No	Nomor Surat	Tanggal	Pengirim	Perihal	Kategori	Berkas	Aksi
1	4213/542/2026	2026-03-11	SMIPA 5 SURABAJA	Pemohonan Mengisi Tawar	Pemohonan	Undut PDF	Detail Edit Hapus
4	100/294/2026	2026-04-16	Dinas PPK Kab Sido	Surat Keterangan Koneksi Jaring Jaringan Fungsional	Surat Keterangan	Undut PDF	Detail Edit Hapus

Gambar 7. Fitur Pencarian Surat

Gambar 7. menunjukkan sistem menyediakan fitur pencarian berdasarkan nomor surat, pengirim, maupun perihal sehingga proses menemukan dokumen arsip menjadi lebih cepat dibandingkan pencarian manual.

B. Hasil Pengujian Algoritma Naïve Bayes

Tabel 1

Hasil Cross Validation					
Fold	Accuracy	Precision	Recall	F1-Score	
	(%)	(%)	(%)	(%)	
1	100	100	100	100	
2	100	100	100	100	
3	100	100	100	100	
4	80	85	80	78,10	
5	100	100	100	100	

Pengujian performa algoritma Naïve Bayes dilakukan menggunakan **5-Fold Cross Validation**[17]. Pada metode ini, seluruh data dibagi menjadi lima bagian (fold). Setiap fold digunakan secara bergantian sebagai data pengujian, sedangkan empat fold lainnya digunakan sebagai data pelatihan. Metode ini bertujuan untuk memperoleh evaluasi model yang lebih objektif karena setiap data memiliki kesempatan yang sama untuk menjadi data uji, sehingga dapat mengurangi bias yang mungkin terjadi apabila hanya menggunakan satu kali pembagian data.

Hasil pengujian menunjukkan bahwa model memperoleh nilai **akurasi rata-rata sebesar 96%, precision sebesar 97,00%, recall sebesar 96,00%, dan F1-Score sebesar 95,62%**. Nilai-nilai tersebut menunjukkan bahwa algoritma Naïve Bayes memiliki kemampuan yang sangat baik dalam mengklasifikasikan surat berdasarkan kategori yang telah ditentukan. Tingginya nilai precision menunjukkan bahwa sebagian besar hasil prediksi yang diberikan model sesuai dengan kategori sebenarnya, sedangkan nilai recall yang tinggi menunjukkan bahwa model mampu mengenali hampir seluruh data pada masing – masing kategori.

Berdasarkan pengujian pada setiap fold, empat fold memperoleh nilai akurasi sebesar **100%**, sedangkan satu fold memperoleh akurasi sebesar **80%**. Perbedaan hasil tersebut menunjukkan bahwa pada salah satu pembagian data terdapat beberapa dokumen yang memiliki karakteristik teks yang mirip antar kategori sehingga menyebabkan terjadinya kesalahan klasifikasi. Meskipun demikian, nilai rata – rata akurasi sebesar 96,00% menunjukkan bahwa performa model tetap stabil dan konsisten dalam melakukan klasifikasi.

Tabel 2

Confusion Matrix			
Aktual / Prediksi	Lainnya	Perizinan	Undangan

Lainnya	6	0	0
Perizinan	0	7	1
Undangan	0	0	10

Hasil **Confusion Matrix**[18] menunjukkan bahwa hanya terdapat **satu kesalahan klasifikasi** dari total **24 data pengujian**. Kesalahan tersebut terjadi pada dokumen kategori **Perizinan** yang diprediksi sebagai kategori **undangan**. Sementara itu, seluruh dokumen pada kategori **Lainnya** dan **Undangan** berhasil di klasifikasi dengan benar. Hal ini menunjukkan bahwa model mampu membedakan karakteristik masing – masing kategori surat dengan tingkat ketepatan yang tinggi.

Secara keseluruhan, hasil pengujian menunjukkan bahwa algoritma Naïve Bayes memiliki performa yang baik dalam melakukan klasifikasi surat pada sistem pengarsipan. Tingginya nilai akurasi precision, recall, dan F1-Score menunjukkan bahwa model mampu memberikan hasil klasifikasi yang andal sehingga layak diterapkan sebagai metode klasifikasi otomatis dalam sistem pengarsipan surat. Implementasi algoritma ini mampu membantu proses pengelompokan dokumen secara cepat, konsisten, dan akurat, sehingga mendukung peningkatan efektivitas pengelolaan arsip pada lingkungan SMP Negeri 5 Sukoharjo.

Hasil penelitian ini menunjukkan bahwa implementasi algoritma Naïve Bayes pada sistem pengarsipan mampu memberikan performa klasifikasi yang tinggi meskipun dibangun tanpa menggunakan library machine learning seperti scikit-learn. Hal tersebut menunjukkan bahwa perhitungan probabilitas secara manual menggunakan Python tetap mampu menghasilkan model klasifikasi yang efektif untuk membantu proses pengelompokan surat secara otomatis. Temuan ini sejalan dengan beberapa penelitian sebelumnya yang menyatakan bahwa algoritma

Naïve Bayes memiliki performa yang baik pada proses klasifikasi dokumen teks [7], [17].

Selain menghasilkan tingkat akurasi yang tinggi, implementasi algoritma Naïve Bayes secara langsung pada aplikasi berbasis web menunjukkan bahwa proses klasifikasi dapat dilakukan secara otomatis tanpa memerlukan perangkat lunak machine learning tambahan. Hal tersebut menjadikan sistem lebih ringan untuk diimplementasikan pada lingkungan sekolah dengan kebutuhan komputasi yang relatif sederhana.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan **Sistem Informasi Pengarsipan Surat Masuk dan Surat Keluar berbasis web** menggunakan **framework Flask** dengan **basis data SQLite** serta menerapkan algoritma **Naïve Bayes** sebagai metode klasifikasi otomatis kategori surat. Sistem yang dikembangkan mampu membantu proses pengelolaan arsip surat, mulai dari penyimpanan data, pencarian arsip, klasifikasi surat secara otomatis, hingga pembuatan laporan dalam format PDF dan Microsoft Excel. Dengan demikian, sistem dapat meningkatkan efektivitas dan efisiensi proses administrasi di SMP Negeri 5 Sukoharjo.

Berdasarkan hasil pengujian menggunakan metode **5-Fold Cross Validation**, algoritma **Naïve Bayes** memperoleh nilai **akurasi rata-rata sebesar 96%**, **precision 97%**, **recall sebesar 96%**, dan **F1-Score sebesar 95,62%**. Selain itu, hasil **Confusion Matrix** menunjukkan bahwa hanya terdapat satu kesalahan klasifikasi dari total 24 data pengujian. Hasil tersebut menunjukkan bahwa algoritma Naïve Bayes memiliki kemampuan yang baik dalam mengklasifikasikan kategori surat sehingga layak diterapkan sebagai metode klasifikasi otomatis pada sistem pengarsipan surat.

Penelitian selanjutnya disarankan menggunakan jumlah data latih yang lebih banyak dan lebih beragam agar performa model klasifikasi dapat ditingkatkan. Selain itu, sistem dapat dikembangkan dengan menambahkan fitur notifikasi surat, disposisi surat, serta implementasi pada platform mobile agar mampu mendukung proses administrasi sekolah secara lebih optimal.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada SMP Negeri 5 Sukoharjo yang telah memberikan izin dan dukungan selama proses pengumpulan data penelitian. Ucapan terima kasih juga disampaikan kepada Program Studi Sistem Informasi Universitas Duta Bangsa Surakarta atas bimbingan, arahan, serta fasilitas yang diberikan sehingga penelitian ini dapat diselesaikan dengan baik.

REFERENSI

- [1] S. Raschka, *Python Machine Learning*, 3rd ed. Birmingham, UK: Packt Publishing, 2019.
- [2] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.
- [3] H. S. Banjarsari, A. Budiman, and P. Farmadi, "Penerapan Algoritma Naïve Bayes pada klasifikasi Data," *Jurnal Ilmiah Informatika*, vol. 5, no. 2, pp. 45-54, 2020.
- [4] A. Saleh, "Implementasi Metode Naïve Bayes dalam Klasifikasi Dokumen," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 3, pp. 551-558, 2020.
- [5] M. A. Hall et al., "The WEKA Data Mining Software: An Update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10-18, 2009.
- [6] M. Grinberg, *Flask Web Development: Developing Web Applications with Python*, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2018.
- [7] R. Elmasri and S.B. Navathe, *Fundamentals of Database Systems*, 7th ed. Boston, MA, USA: Pearson, 2016.
- [8] S. Silberschatz, H. F. Korth, and S. Sudarshan, *Database System Concepts*, 7th ed. New York, NY, USA: McGraw-Hill, 2020.
- [9] D. C. Montgomery, *Design and Analysis of Experiments*, 10th ed. Hoboken, NJ, USA: John Wiley & Sons, 2019.
- [10] A. Dennis, B. H. Wixom, and D. Tegarden, *Systems Analysis and Design: An Object-Oriented Approach with UML*, 6th ed. Hoboken, NJ, USA: John Wiley & Sons, 2021.
- [11] R. S. Pressman and B. R. Maxim, *Software Engineering: A Practitioner's Approach*, 9th ed. New York, NY, USA: McGraw-Hill Education, 2020.