

KLASIFIKASI SURAT DIGITAL MENGGUNAKAN ALGORITMA NAÏVE BAYES

Ilhamsah Kasim¹, Yuyun², Supriadi Sahibu³

¹Kanwil Kementerian Agama Propinsi Sulawesi Selatan, ^{2,3}STMIK Handayani Makassar

¹baso.cengho@gmail.com, ²yuyunwabula@handayani.ac.id, ³supriadi@handayani.ac.id

ABSTRAK

Tujuan penelitian ini adalah untuk mengklasifikasi surat masuk pada kantor kementerian agama propinsi Sulawesi selatan. Dalam mengklasifikasi surat ini, kami menggunakan algoritma Naïve Bayes dan TF-IDF. Dalam penelitian ini jumlah sampel data corpus surat 1036 record, yang dibagi dalam 80% training dan 20% testing. Sampel data training algoritma Naïve Bayes di implementasikan dengan menghitung nilai bobot dari class tertinggi berdasarkan data training dengan data testing sehingga menghasilkan probabilitas tertinggi. Hasil pengolahan data mendapatkan nilai *Correctly Classified Instance* sebesar 86.245799% dan *Incoreectly Classified Instance* sebesar 13.754200% serta hasil pengujian dengan menggunakan confusion matrix mendapatkan nilai precision sebesar 86%, Recall 86 % dan Akurasi sebesar 76%.

Kata kunci: klasifikasi, naive, Bayes, surat, digital

1. PENDAHULUAN

Perkembangan teknologi informasi telah berpengaruh pada pengelolaan data. Bagian Tata Usaha Kantor Wilayah Kementrian Agama Provinsi Sulawesi Selatan memiliki tanggung jawab dalam mengelola surat masuk. Surat-surat yang dikelola oleh tata usaha dikategorikan menjadi dua jenis yaitu jenis surat undangan dan jenis surat umum. Setiap bulan bagian tata usaha menerima surat masuk sebanyak 1.800. Surat-surat tersebut kemudian disimpan sebagai arsip dan digunakan kembali pada saat dibutuhkan. Proses pengerjaan surat masuk masih dikerjakan secara manual. Dari hasil analisis awal, sebelum surat masuk diberikan kepada penerima, bagian tatauusaha membaca dan menganalisis tujuan, perihal dan kepada siapa surat diberikan. Permasalahan yaitu sulit dalam monitoring dan tracking surat, kadang surat hilang tidak terlacak, membutuhkan ruang fisik untuk menyimpan surat yang secara terus menerus ada. Hal ini tentu mengganggu efektivitas kinerja organisasi Untuk mengatasi hal tersebut perlu pengelolaan surat yang dikelola secara digital yaitu sistem yang secara otomatis dapat mengenali informasi surat masuk. Dalam penelitian ini, sebanyak 1036 dokumen surat dianalisis.

Penelitian tentang pengelolaan surat digital telah dilakukan oleh beberapa peneliti, misalnya [1] yaitu klasifikasi Surat Menggunakan Metode Naïve Bayes Pada Sistem Informasi Manajemen Surat” dimana Sistem ini dibuat dengan metode *Naïve Bayes Classification* (NBC). Sistem monitoring surat menggunakan metode Naïve Bayes dibangun berbasis web dengan bahasa pemrograman PHP. Data yang diolah berekstensi (.pdf) kemudian dikonversi dalam bentuk teks menggunakan OCR. Penelitian ini menghasilkan akurasi yang cukup rendah yaitu 33%. Dalam penelitian lain [2] mengimplementasi algoritma multinomial naive bayes classifier pada sistem klasifikasi surat keluar. Peneliti mengkaji Multinomial Naive Bayes untuk mengklasifikasi surat keluar. Sistem klasifikasi didukung dengan confix-stripping stemmer menggunakan model TF-IDF untuk pembobotan kata dan diuji menggunakan *confusion matrix*. Pengujian mereka menghasilkan tingkat *accuracy*, *precision*, *recall*, dan *F-measure* masing-masing 89,58%, 79,17%, 78,72%, dan 77,05%. Kemudian [3] mengusulkan algoritma boyer-moore pada klasifikasi kode surat berbasis android di kantor kecamatan. Aplikasi diuji dengan menggunakan metode black-box testing. Berdasarkan hasil black-box testing dan pengujian penilaian yang dilakukan pengguna, aplikasi dapat berjalan secara lancar dan dinilai efektif serta layak untuk digunakan dalam mencari klasifikasi kode surat. Dalam penelitian lain [4] mengusulkan Algoritma Decision Tree untuk Klasifikasi Surat pada Surat pada dinas Komunikasi dan Informatika kota Kediri berbasis Android. menggunakan Android. Aplikasi ini ternyata mampu menampilkan hasil dari buka dokumen, dokumen baru, penggabungan dokumen, dan pencarian dokumen pada Android [4]. Ada juga aplikasi penyimpanan *file*



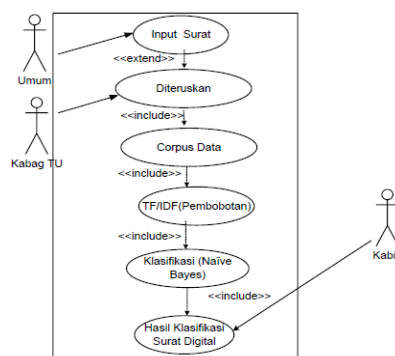
Berdasarkan hasil penelitian dengan menggunakan pengujian fungsional yaitu metode blackbox testing menunjukkan bahwa sistem mampu berjalan dengan baik dan berhasil menampilkan surat sesuai dengan prioritas surat yang mengacu sifat dan kategori surat. Kemudian untuk algoritma decision tree yang diimplementasikan pada backend aplikasi berjalan dengan baik dan sesuai dengan hasil yang diinginkan. Kemudian [5] mengusulkan Algoritma Fuzzy K-Nearest Neighbour untuk ekstraksi informasi pada dokumen surat masuk. Hasil pengujian akurasi menunjukkan bahwa nilai akurasi klasifikasi Fuzzy K-NN lebih baik daripada klasifikasi K-NN dengan 87,52 % berbanding 85,17%. Berdasarkan hasil tersebut, disimpulkan algoritma Fuzzy K-NN dapat diterapkan menjadi klasifikasi dalam ekstraksi informasi dokumen surat masuk. Dalam penelitian ini kami mengusulkan algoritma naïve bayes untuk klasifikasi dan algoritma TF-IDF untuk pembobotan masing-masing dokumen surat.

Dari hasil literatur review, maka tujuan penelitian ini adalah mengklasifikasi surat masuk berupa surat undangan dan surat umum berdasarkan kriteria atau parameter yang digunakan dari masing-masing surat. Model yang diusulkan menggunakan model mesin learning dengan algoritma naïve bayes. Algoritma ini adalah jenis metode yang menggunakan model probabilitas dan statistik untuk memprediksi peluang dimasa depan berdasarkan pengalaman di masa sebelumnya [6]. Alasan menggunakan model ini adalah handal, memiliki kinerja baik dan sanggup mengklasifikasi jumlah data yang besar [7][8][9][10][11]. Dalam penelitian ini model yang diusulkan untuk mengklasifikasikan surat pada kantor Kementerian Agama Propinsi Sulawesi Selatan.

2. METODE PENELITIAN

Metode pengembangan sistem yang digunakan dalam penelitian ini adalah model waterfall. Model ini melakukan pendekatan secara sistematis dan berurutan dalam membangun sistem. Tahapan dari metode waterfall yang pertama adalah tahap analisis yaitu menganalisa kebutuhan user, analisa perangkat lunak dan perangkat keras yang dibutuhkan dalam pengembangan sistem. Tahapan kedua adalah desain sistem. Tujuan dari tahap ini adalah memberikan gambaran tentang apa yang akan dikerjakan dan bagaimana tampilannya. Tahap ini memenuhi semua kebutuhan pengguna sesuai dengan hasil yang dianalisa seperti rancangan tampilan pengembangan sistem dan membantu mendefinisikan arsitektur sistem secara keseluruhan. Tahap ketiga yaitu coding (Penulisan Kode Program). Aktivitas pada tahap ini dilakukan pengkodean sistem. Penulisan kode program merupakan tahap penerjemahan desain sistem yang telah dibuat ke dalam bentuk perintah-perintah yang dimengerti komputer dengan mempergunakan bahasa pemrograman. Tahapan ini merupakan tahapan secara nyata dalam mengerjakan suatu sistem. Sistem ini bahasa pemrograman yang dipakai adalah PHP dan database MySQL. Tahap ke Empat yaitu Testing (Pengujian Program). Pengujian dilakukan untuk memastikan bahwa software yang dibuat telah sesuai dengan desainnya dan semua fungsi dapat dipergunakan dengan baik tanpa ada kesalahan. Tahap ke Lima yaitu Operation and Maintenance (Pemeliharaan Program). Tahap ini merupakan tahap terakhir dalam metode waterfall. Sistem dapat di implementasikan. Pemeliharaan mencakup koreksi dari berbagai error yang tidak ditemukan pada tahap-tahap terdahulu, perbaikan atas implementasi dan pengembangan unit sistem, serta pemeliharaan program. Pemeliharaan sistem dapat dilakukan oleh seorang administrator untuk meningkatkan kualitas sistem agar jauh lebih baik.

Untuk mengimplementasikan model pada sistem, kami menggunakan algoritma klasifikasi Naïve Bayes. Proses klasifikasi dilakukan melalui 2 tahap, di antaranya tahap training dan testing. Tahapan pembelajaran atau training merupakan proses mengekstraksi surat yang telah diketahui kategorinya. Proses ini dilakukan dengan penentuan dan pembentukan bag of word dengan model TF-IDF pada setiap dokumen training. Setiap kata yang disimpan dianalisis nilai weighting terhadap semua dokumen pada setiap klasifikasi sehingga dapat menghitung probabilitas kata yang nantinya akan digunakan pada tahap testing. Berdasarkan Analisis sistem yang sedang berjalan maka penulis memberikan satu solusi pemecahan masalah dengan perancangan, klasifikasi surat digital dengan algoritma *machine Learning* sebagai berikut:



Gambar 2.1 Use Case Diagram model yang diusulkan

Gambar diatas merupakan diagram usecase yan diusulkan dimana perancangan ini terdapat tiga aktor yaitu Umum, kabag TU dan Kabid. Pada bagian Umum menginput surat, kemudian surat dari bagian kabag TU akan diteruskan dengan mengaktifkan fitur diteruskan yang terhubung dengan corpus data, selanjutnya melakukan pembobotan dengan TF/IDF. Setelah melakukan pembobotan proses klasifikasi dengan algoritma naïve bayes kemudian ditampilkan hasil klasifikasi digital yang dapat dilihat oleh kabid bagian masing-masing. Dalam penelitian ini, sistem akan menggunakan algoritma naïve bayes. Untuk pengujian diukur menggunakan confusion matrix dengan memperhitungkan accuracy, recall, precision, dan F-measure pada setiap klasifikasi.

3. HASIL DAN PEMBAHASAN

A. Hasil

Dalam klasifikasi surat masuk ini pertama-tama adalah menentukan data corpus surat. Kemudian melakukan pembobotan dengan menyiapkan variable penampung data inputan. Dalam pembobotan untuk text masing-masing surat menggunakan algoritma TF-IDF. Dilanjutkan dengan membuat data input untuk simulasi dengan fungsi random untuk masing-masing atribut.

TABEL 2 : Jumlah Data Kelas DISPOSISI KE

Disposisi Ke	Jumlah
KABAG TATA USAHA	1511
KABID PENDIDIKAN AGAMA ISLAM	251
KABID PENDIDIKAN DINIYAH DAN PONDOK PESANTREN	252
KABID PENDIDIKAN MADRASAH	2182
KABID PENERANGAN AGAMA ISLAM, ZAKAT DAN WAKAF	1403
KABID PENYELENGGARAAN HAJI DAN UMRAH	2330
KABID URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH	2751
KABID URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH, KRISTEN, KATOLIK	16
KABID URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH/KATOLIK	12
KEPALA KANTOR	1396
PENDIDIKAN MADRASAH	35
PENYELENGGARAAN HAJI DAN UMRAH	160
TATA USAHA	18
URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH	179

Gambar 3.1 Daftar Kelas untuk Klasifikasi Surat

NO	DARI	PERIHAL	DITUJUKAN	URAIAN/ISI SURTI	DISPOSISI KE	
1	1	KUA Kec. Manggala, Kota Makassar	Undangan peresmian Gedung KUA	KABID URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH	dijadwalkan	URUSAN AGAMA ISLA
2	2	KUA Kec. Manggala, Kota Makassar	Undangan peresmian Gedung KUA	KEPALA KANTOR	dijadwalkan	KEPALA KANTOR
3	3	KUA Kec. Manggala, Kota Makassar	Undangan peresmian Gedung KUA	KABAG TATA USAHA	dijadwalkan	KABAG TATA USAHA
4	4	Madrasah Aliyah DDI Mangkoso	Proposal Bantuan Rehab Madrasah	KEPALA KANTOR	ditindaklanjuti	KABID PENDIDIKAN M
5	5	Travel Umrah Tazkiyah	Undangan Pelepasan Jamaah Umrah	KABID PENYELENGGARAAN HAJI DAN UMRAH	dijadwalkan	KABID PENYELENGGA
6	6	Madrasah Tsanawiyah Negeri 1 Makassar	Undangan Peresmian Gedung Kelas Baru	KEPALA KANTOR	Diwakili	KABID PENDIDIKAN M
7	7	Madrasah Tsanawiyah Negeri 1 Makassar	Undangan Peresmian Gedung Kelas Baru	KABAG TATA USAHA	dijadwalkan	KABAG TATA USAHA
8	8	Madrasah Tsanawiyah Negeri 1 Makassar	Undangan Peresmian Gedung Kelas Baru	KABID PENDIDIKAN MADRASAH	dijadwalkan	KABID PENDIDIKAN M
9	9	Madrasah Aliyah DDI Jampue	Proposal Bantuan Rehab Madrasah	KEPALA KANTOR	ditindaklanjuti	KABID PENDIDIKAN M
10	10	Madrasah Tsanawiyah DDI Jampue	Proposal Bantuan Rehab Madrasah	KEPALA KANTOR	ditindaklanjuti	KABID PENDIDIKAN M
11	11	Madrasah Ibtidaiyah DDI Jampue	Proposal Bantuan Rehab Madrasah	KEPALA KANTOR	ditindaklanjuti	KABID PENDIDIKAN M
12	12	Kantor Kemenag Kab. Bone	Usulan Kenaikan Pangkat	KEPALA KANTOR	ditindaklanjuti	KABAG TATA USAHA
13	13	Kantor Kemenag Kab. Soppeng	Usulan Kenaikan Pangkat	KEPALA KANTOR	ditindaklanjuti	KABAG TATA USAHA
14	14	Kantor Kemenag Kab. Wajo	Usulan Kenaikan Pangkat	KEPALA KANTOR	ditindaklanjuti	KABAG TATA USAHA
15	15	Sekretariat Jendral Kementerian Agama RI	Edaran Pemetaan Kelas Klasifikasi Jabatan	KEPALA KANTOR	ditindaklanjuti	KABAG TATA USAHA
16	16	Sekretariat Jendral Kementerian Agama RI	Pelaksanaan Rekrutmen CPNS Kemenag	KEPALA KANTOR	ditindaklanjuti	KABAG TATA USAHA
17	17	Madrasah Aliyah Negeri Insancendekia	Undangan Peresmian Gedung Kelas Baru	KEPALA KANTOR	Diwakili	KABID PENDIDIKAN M
18	18	Madrasah Aliyah Negeri Insancendekia	Undangan Peresmian Gedung Kelas Baru	KABAG TATA USAHA	dijadwalkan	KABAG TATA USAHA
19	19	Madrasah Aliyah Negeri Insancendekia	Undangan Peresmian Gedung Kelas Baru	KABID PENDIDIKAN MADRASAH	dijadwalkan	KABID PENDIDIKAN M
20	20	Madrasah Aliyah Negeri Insancendekia	Proposal Pembangunan Gedung Kelas Baru Melalui SBSN	KEPALA KANTOR	ditindaklanjuti	KABID PENDIDIKAN M

Gambar 3.2 Daftar Corpus Surat

B. Pembahasan

Dalam penelitian ini jumlah sampel data corpus surat 1036 record, yang dibagi dalam 80% training dan 20% testing. Adapun proses testing yaitu

- Indexing terhadap data testing memiliki proses yang sama dengan indexing pada data training. Nilai term frequency tersebut akan digunakan untuk memangkatkan nilai conditional probabilities dari model.
- Proses matching adalah proses untuk menemukan term yang sama – sama muncul pada model maupun pada data testing
- Hasil matching tersebut kemudian digabungkan dengan nilai conditional probabilitiesnya

TABEL 5: Probabilitas Atribut Atribut DITUJUKAN

DITUJUKAN	DISPOSISI KE							
	KABAG TATA USAHA	KABID PENDIDIKAN AGAMA ISLAM	KABID PENDIDIKAN DINIYAH DAN PONDOK PESANTREN	KABID PENDIDIKAN MADRASAH	KABID PEMERANGAN AGAMA ISLAM, ZAKAT DAN WAKAF	KABID PENYELENGGARAAN HAJI DAN UMRAH	KABID URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH	KABID URUS/AGAMA ISLAM PEMBINAAN SYARIAH, KRI KATOLIK
agama	7.9923273657289E-5	0.0022403584578532	7.9923273657289E-5	7.9923273657289E-5	0.0086611137782045	7.9923273657289E-5	0.010321651464234	7.9923273657
dan	0.0039005760921747	7.9923273657289E-5	0.0081613058089294	7.9923273657289E-5	0.038726196191391	0.03664586336134	0.084253480556889	0.0006401024
diniyah	7.9923273657289E-5	7.9923273657289E-5	0.00096015362457993	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	7.9923273657
haji	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	0.03664586336134	7.9923273657289E-5	7.9923273657
islam	7.9923273657289E-5	0.0022403584578532	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	0.010321651464234	7.9923273657
islam,	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	0.0086611137782045	7.9923273657289E-5	7.9923273657289E-5	7.9923273657
kabag	0.036165786525944	7.9923273657289E-5	7.9923273657289E-5	0.00048007681228097	7.9923273657289E-5	7.9923273657289E-5	7.9923273657289E-5	7.9923273657

Gambar 3.3 Conditional Probabilitas Kelas

Untuk mendapatkan nilai Probabilitas mengalikan nilai prior probabilities dengan total nilai conditional probabilities untuk masing – masing kelas dapat dilihat pada gambar berikut

Correctly Classified Instance - 86.245799327892%
 Incorrectly Classified Instance - 13.754200672108%

DARI-Travel Umrah Sisi Tour yang memiliki kata **travel** sbg bobot tertinggi dengan nilai bobot **462.459**
PERIHAL-Undangan Pelepasan Jamaah Umrah yang memiliki kata **undangan** sbg bobot tertinggi dengan nilai bobot **478.995**
DITUJUKAN-KABID PENYELENGGARAAN HAJI DAN UMRAH yang memiliki kata **penyelenggaraan** sbg bobot tertinggi dengan nilai bobot **1323.79**
URAIAN-dijadwalkan yang memiliki kata **dijadwalkan** sbg bobot tertinggi dengan nilai bobot **1929.72**

Masuk ke kelas DISPOSISI KE - **KABID PENYELENGGARAAN HAJI DAN UMRAH** dengan probabilitas sebesar **7.7798697488804E-8**

Gambar 3.4 Hasil Probabilitas Kelas

Setelah melalui proses tersebut dilanjutkan dengan proses pembobotan TF-IDF dilanjutkan proses kalsifikasi Gambar berikut:

Rekomendasi system dari inputan di atas, surat didisposisi ke bidang **KABID PENDIDIKAN MADRASAH**

PEMBOBOTAN TF-IDF

Id	Term	Atribut	Class	Count	Bobot
1	asosiasi	dari	KABID PENDIDIKAN AGAMA ISLAM	18	59.3251
2	guru	dari	KABID PENDIDIKAN AGAMA ISLAM	18	46.8484
3	guru	dari	KABAG TATA USAHA	3	7.80807
4	pai	dari	KABID PENDIDIKAN AGAMA ISLAM	18	59.3251
5	indonesi	dari	KABID PENDIDIKAN AGAMA ISLAM	18	59.3251
6	(agpaai)	dari	KABID PENDIDIKAN AGAMA ISLAM	18	59.3251
7	undangan	perihal	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH	56	45.4121
8	undangan	perihal	KEPALA KANTOR	262	212.464
9	undangan	perihal	KABAG TATA USAHA	436	353.566
10	undangan	perihal	KABID PENYELENGGARAAN HAJI DAN UMRAH	166	134.614
11	undangan	perihal	KABID PENDIDIKAN MADRASAH	143	115.963

Gambar 3.5. Proses pembentukan bobot masing-masing kata dengan TF-IDF

81	Islam	ditujukan	KABID URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH, KRISTEN, KAT	1	0.730888
----	-------	-----------	--	---	----------

SARAN DISPOSISI
Naive Bayes - Persuaratan

TABEL 1 - Data Training

No	DARI	PERIHAL	DITUJUKAN	URAIAN	DISPOSISI KE
1	kua	undangan	kabid	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
2	kua	undangan	urusan	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
3	kua	undangan	agama	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
4	kua	undangan	Islam	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
5	kua	undangan	dan	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
6	kua	undangan	pembinaan	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
7	kua	undangan	syariah	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
8	kua	peresmian	kabid	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
9	kua	peresmian	agama	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
10	kua	peresmian	Islam	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
11	kua	gedung	kabid	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH
12	kua	gedung	agama	dijadwalkan	URUSAN AGAMA ISLAM DAN PEMBINAAN SYARIAH

Gambar 3.6. Proses klasifikasi

Setelah melalui proses klasifikasi dilanjutkan dengan dengan pengujian hasil klasifikasi. Untuk melakukan pengujian ini kami menggunakan *confusion matrix*.

Tabel 3.1 Confution matriks

Predicate Class			
		Positive	Negative
Actual Class	Positive	(TP)=74	(FN)=12
	Negative	(FP)=12	(TN)=2

True positives adalah jumlah record positif yang diklasifikasikan sebagai positif, false positives adalah jumlah record positif yang diklasifikasikan sebagai negatif, false negatives adalah jumlah record negatif yang diklasifikasikan sebagai positif, true negatives adalah jumlah record negatif yang diklasifikasikan sebagai negatif. Data uji yang dimasukkan ke dalam confusion matrix, akan dihitung nilai-nilai recall, precision dan accuracy.

Tabel 3.2 Pengujian dengan Confution matriks

Menentukan Nilai <i>Precision</i> (<i>p</i>)	Menentukan Nilai <i>Recall</i> (<i>r</i>)	Menentukan Nilai Akurasi
$P = \frac{TP}{TP + FP} \quad P = \frac{74}{74 + 12}$ <p><i>Precision</i> (<i>p</i>) = 0,86 * 100 % = 86%</p>	$Recall = \frac{TP}{TP + FN} \quad Recall = \frac{74}{74 + 12}$ <p><i>Recall</i> (<i>r</i>) = 0,86 * 100 % = 86%</p>	$akurasy = \frac{TP + TN}{TP + TN + FN + FP}$ $akurasy = \frac{74 + 2}{74 + 2 + 12 + 12}$ <p>akurasy = 0,76 * 100 % = 76 %</p>

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan Hasil Penelitian dalam penelitian ini telah dapat di implementasikan Algoritma Naïve Bayes sebagai Klasifikasi Surat Digital Menggunakan Algoritma Machine learning dan kesimpulan yang dapat diambil dari pembangunan sistem ini adalah:

1. Algoritma Naïve Bayes digunakan untuk klasifikasi surat berdasarkan data surat masuk baru berdasarkan sampel dari data training. Hasil Klasifikasi dari sampel data training Algoritma Naïve Bayes telah dapat di implementasikan dengan menghitung nilai bobot dari class tertinggi berdasarkan data training dengan data testing sehingga menghasilkan model atau hasil probabilitas tertinggi yang dapat di teruskan berdasarkan tujuan surat
2. Hasil pengolahan data dengan algoritma naïve bayes menghasilkan nilai untuk sampel pengolahan data class Kabag Tata Usaha mendapatkan nilai Correctly Classified Instance sebesar 86.245799% dan Incorectly Classified Instance sebesar 13.754200%.



4.2 Saran

Berdasarkan kesimpulan yang telah dijabarkan, maka dapat disarankan yaitu:

Pada halaman Edit foto *scanner* ditampilkan halaman untuk mengubah foto menjadi format hasil *scan*. Tampilan dari halaman Edit foto *scanner* ditunjukkan pada Gambar 3.6.

1. Penelitian ini dapat dikembangkan dengan mengimplementasikan algoritma lain, untuk memperoleh akurasi terbaik
2. Menggabungkan sistem klasifikasi yang otomatis dan manual. Hal ini dapat memberikan user kemampuan untuk memperbaiki kesalahan klasifikasi oleh sistem. Sehingga data yang dikelola akan menjadi semakin baik

DAFTAR PUSTAKA

- [1] M. D. Trisetiyo and J. S. Wibowo, "Klasifikasi Surat Menggunakan Metode Naïve Bayes Pada Sistem Informasi Manajemen Surat," *Pros. SENDI*, pp. 978–979, 2019.
- [2] A. H. Setianingrum, D. H. Kalokasari, and I. M. Shofi, "Implementasi Algoritma Multinomial Naive Bayes Classifier," *J. Tek. Inform.*, vol. 10, no. 2, pp. 109–118, 2018, doi: 10.15408/jti.v10i2.6822.
- [3] Y. Muharam and Y. Bahtiar, "Pembangunan Aplikasi Klasifikasi Kode Surat Berbasis Android Menggunakan Algoritma Boyer-Moore Di Kantor Kecamatan Ciparay," *J. Inform.*, vol. 08, no. 4, pp. 50–57, 2021.
- [4] M. N. Baharudin and I. K. Dwi Nuryana, "Implementasi Algoritma Decision Tree untuk Klasifikasi Surat pada Aplikasi Mobile E-Surat Dinas Komunikasi dan Informatika Kota Kediri Berbasis Android," *J. Informatics Comput. Sci.*, vol. 4, no. 01, pp. 76–85, 2022, doi: 10.26740/jinacs.v4n01.p76-85.
- [5] A. Sulaiman, "Ekstraksi Informasi Pada Dokumen Surat Masuk Menggunakan Algoritma Fuzzy K-Nearest Neighbour (Fuzzy K-NN)," pp. 1–8, 2018, [Online]. Available: <https://elibrary.unikom.ac.id/id/eprint/483/>
- [6] M. F. Rifai, H. Jatnika, and B. Valentino, "Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS)," *Petir*, vol. 12, no. 2, pp. 131–144, 2019, doi: 10.33322/petir.v12i2.471.
- [7] A. Achmad, "Klasifikasi Pembibitan Udang Vanamey Yang Ideal Menggunakan Algoritma Naive Bayes," vol. 3, no. 3, pp. 161–168, 2022.
- [8] M. Lutfi, H. Muttaqien, A. Apriliani, H. Zainuddin, and Y. Yuyun, "Application of the Naïve Bayes Algorithm and Simple Exponential Smoothing for Food Commodity Prices Forecasting," 2019, doi: 10.4108/eai.2-5-2019.2284613.
- [9] S. Samasil, Y. Yuyun, and H. Hazriani, "Klasifikasi Mahasiswa Berpotensi Drop Out Menggunakan Algoritma Naive Bayes Dan Decision Tree," *J. Ilm. Ilmu Komput.*, vol. 8, no. 2, pp. 108–114, 2022, doi: 10.35329/jiik.v8i2.242.
- [10] S. Fitri, "Perbandingan Kinerja Algoritma Klasifikasi Naïve Bayesian," *Data Manaj. dan Teknol. Inf.*, vol. 15, no. 1, pp. 33–37, 2014, [Online]. Available: <http://ojs.amikom.ac.id/index.php/dasi/article/view/180%0Ahttps://ojs.amikom.ac.id/index.php/dasi/article/viewFile/180/165>
- [11] G. W. N. Wibowo and M. A. Manan, "Penerapan Algoritma Naive Bayes Untuk Prediksi Heregistrasi Calon Mahasiswa Baru," *JTINFO J. Tek.*, vol. 1, no. 1, pp. 1–10, 2022, [Online]. Available: <https://journal.unisnu.ac.id/JTINFO/article/view/126>

