PERBANDINGAN *ALGORITME-ALGORITME* PEMBELAJARAN MESIN PADA KLASIFIKASI SMS *SPAM*

Edi Zuviyanto [1)], Teguh Bharata Adji[2)], Noor Akhmad Setiawan [3)]

[1),2),3)]Departemen Teknik Elektro dan Teknologi Informasi FT UGM
Jl. Grafika No.2 Kampus UGM
Email : edi.ti14@mail.ugm.ac.id

Abstrak. *Short Message Service (SMS) merupakan salah satu media komunikasi yang termurah. Dengan murahnya tarif layanan SMS, membuat pihak-pihak yang tidak bertanggungjawab menggunakannya untuk kepentingan bisnis dan tindak kejahatan. Hal ini, menyebabkan kerugian bagi masyarakat yang menggunakan layanan SMS. Oleh karena itu, diperlukan teknik klasifikasi teks untuk membedakan pesan SMS yang berisi spam dan bukan spam (ham). Beberapa penelitian sudah dilakukan untuk mengklasifikasikan SMS spam dan ham. Namun, belum ditemukan algoritme klasifikasi yang baik untuk klasifikasi SMS spam. Penelitian ini bertujuan untuk mencari algoritme klasifikasi yang baik untuk klasifikasi SMS spam dengan membandingkan beberapa algoritme klasifikasi machine learning. Beberapa algoritme tersebut yaitu C4.5, KNN, NBC, dan SVM. Penelitian dilakukan dengan beberapa langkah yaitu pengumpulan data, preprocessing, pembobotan fitur, klasifikasi, evaluasi dan perbandingan akurasi. Hasilnya, pengujian klasifikasi SMS dengan menggunakan algoritme SVM memiliki tingkat akurasi lebih baik daripada menggunakan algoritme C4.5, KNN maupun NBC. Nilai akurasi yang diperoleh dengan menggunakan metode SVM sebesar 94,60%, sedangkan nilai akurasi dengan menggunakan metode C4.5, KNN dan NBC berturut-turut sebesar 85,86%, 77,50% dan 86,10 %.*

Kata kunci : *Klasifikasi, SMS spam, Naive bayes classifier, Support vector machine, C4.5, KNN.*

1. Pendahuluan

Short Message Service (SMS) merupakan salah satu media komunikasi yang masih banyak digunakan, karena mekanisme penggunaannya sangat mudah. Selain itu, tarif dari layanan SMS sangat murah dan masih basnyak bonus yang diberikan kepada pengguna [1]. Namun, adanya tarif SMS yang murah dan banyaknya bonus SMS menyebabkan adanya pihak-pihak yang tidak bertanggungjawab untuk melakukan tindak kejahatan seperti SMS penipuan. SMS tersebut termasuk dalam salah satu jenis SMS spam.

SMS spam merupakan SMS yang tidak diinginkan oleh penerima [2]. Biasanya informasi yang terdapat pada SMS spam bertujuan untuk pemasaran, promosi, pengiklanan, penipuan dan lain-lain [3]. Hal itu, menyebabkan kerugian bagi masyarakat pengguna layanan SMS. Salah satu yang bisa dilakukan untuk mengatasi SMS spam adalah dengan melakukan klasifikasi SMS spam dan non spam (*ham*). Klasifikasi ini dilakukan untuk membedakan pesan yang berisi spam dan *ham*. Sehingga pengguna *handphone* tidak dirugikan dengan adanya SMS spam.

Beberapa penelitian sudah dilakukan untuk mengatasi masalah SMS spam diantaranya yaitu penelitian yang dilakukan oleh Dewi [4] , dimana algoritme yang digunakan untuk klasifikasi SMS spam dan *ham* adalah *naïve bayes classifier* (NBC). NBC merupakan salah satu algoritme yang efektif digunakan untuk klasifikasi teks dengan jumlah yang besar. Dalam penelitian [4] menggunakan *RapidMinner* untuk mencari konfigurasi terbaik NBC untuk mencapai tingkat akurasi yang maksimal. Kemudian, NBC juga digunakan oleh Indiarto [5] untuk mengklasifikasi SMS spam dan *ham* dengan *term frequency*(TF) sebagai metode pembobotan fitur. Hasilnya, NBC dapat memberikan tingkat akurasi yang cukup untuk menyaring SMS yang masuk.

Selanjutnya, penelitian tentang SMS spam juga dilakukan oleh Apandi [6] dengan menggunakan algoritme *Support Vector Machine* (SVM) untuk klasifikasi SMS spam dan *ham*. Penelitian ini menggunakan metode evaluasi yaitu *K-fold Cross Validation* (*K-fold*). Beberapa referensi menyebutkan bahwa SVM akan memberikan hasil ketepatan klasifikasi yang lebih baik bila dikombinasikan dengan teknik partisi data *K-fold Cross Validation* [7].

Ada pun beberapa penelitian tentang klasifikasi teks yang mempunyai hasil performa yang tidak kalah bagus dengan SVM dan NBC yaitu Gata [8] mengusulkan algoritme *K-Nearest Neighbor* (KNN) untuk mangklasifikasikan SMS pada aplikasi SMS *gateway* yang digunakan oleh LKBN ANTARA. Hasil nilai akurasi yang didapat sangat baik yaitu sebesar 96,15%. Kemudian, penelitian [9] tentang perbandingan performa dari beberapa algoritme antara lain svm, *neural network*, *naïve bayes* dan C4.5 untuk mengklasifikasikan spam email. Hasil terbaik diperoleh saat menggunakan algoritme C4.5 untuk klasifikasi spam email.

Dari semua hasil penelitian yang sudah dilakukan, belum ditemukan algoritme klasifikasi yang tepat untuk klasifikasi SMS spam. Oleh karena itu, dilakukan perbandingan tingkat akurasi dari beberapa algoritme klasifikasi untuk mengklasifikasikan SMS ke dalam kelas SMS spam dan *ham*. Beberapa algoritme klasifikasi yang digunakan yaitu C4.5, KNN, NBC dan SVM.

2. Metode

Metode penelitian yang diusulkan pada penelitian ini yaitu pengumpulan data, *preprocessing*, pembobotan fitur, klasifikasi, dan evaluasi. *Preprocessing* masih memiliki sub-proses lagi yaitu *case folding*, normalisasi kalimat, *stopword removal*, dan *stemming*. Berikut penjelasan beberapa proses dan sub-proses pada metode penelitian yang diusulkan.

2.1 Pengumpulan Data

Penelitian ini menggunakan sebagian besar data dari penelitian Mair [10] dan sebagian lagi mengambil data SMS yang baru dari *handphone*. Data SMS yang digunakan sebanyak 100 SMS yang terdiri dari 50 SMS spam dan 50 SMS *ham*. Data tersebut masih belum tersetruktur sesuai kebutuhan proses klasifikasi. Oleh karena itu, perlu dilakukan pemrosesan teks atau *preprocessing*.

2.2 Preprocessing

Pemrosesan teks (*Preprocessing*) bertujuan untuk mengubah data yang belum terstruktur menjadi data yang tersetruktur sesuai kebutuhan proses mining. *Preprocessing* yang dilakukan, diantaranya adalah :

- ***Case Folding***

Case folding yaitu proses untuk mengubah huruf dari isi SMS menjadi huruf kecil semua, sehingga SMS tidak ada yang memiliki huruf kapital. Pada penelitian ini semua huruf dirubah ke dalam huruf kecil.

- **Normalisasi Kalimat**

Normalisasi kalimat merupakan proses untuk menormalkan kalimat yang gaul menjadi normal, sehingga kalimat dapat dipahami dan sesuai dengan KBBI [11] . Data SMS yang digunakan masih banyak kalimat atau kata-kata gaul yang masih sulit dipahami. Hasil dari normalisasi kalimat akan terlihat beberapa kata gaul yang memiliki makna yang sama dengan yang normal.

- ***Stopword Removal***

Proses ini berfungsi untuk menghilangkan kata-kata yang tidak mempengaruhi secara signifikan dalam proses klasifikasi. *Stopword* dapat berupa kata depan, kata penghubung, dan kata pengganti [12] . Misalnya kata yang, ini, dari, dengan, ke, di, dan sebagainya.

- ***Stemming***

Proses ini berfungsi untuk mengubah kata yang berimbuhan menjadi sebuah kata dasar. Pada penelitian ini, semua kata pada isi SMS akan diubah kedalam kata dasar.

2.3 Pembobotan Fitur

Penelitian ini menggunakan metode pembobotan fitur *Term Frequency – Inverse Document Frequency* (TF-IDF). TF-IDF adalah metode pembobotan fitur yang menghitung bobot dari kata atau istilah (term) yang muncul dalam sebuah dokumen. Nilai bobot kata pada metode TF-IDF didapat dari hasil perkalian antara nilai TF dengan nilai IDF. Proses pembobotan dari *term* yang ada menggunakan persamaan sebagai berikut :

$$w_{t.d} = \mathrm{tf}_{t,d}.\log\frac{n}{\mathrm{dfi}} \quad \text{.....(1)}$$

dimana,
$w_{t.d}$ = frekuensi *term t* pada dokumen *d.*
$log\frac{n}{dfi}$ = *Inverse document frequency* (*idf*)
n = banyaknya dokumen
dfi = banyaknya dokumen yang memiliki *term t.*

3 Klasifikasi

Klasifikasi adalah sebuah proses untuk mencari model atau fungsi yang menjelaskan dan membedakan kelas dari data dengan tujuan untuk melakukan prediksi dari kelas suatu objek dimana tidak diketahui label dari kelas tersebut. Model ini berasal dari analisis dari kumpulan training data atau objek data dimana kelas dari label sudah diketahui. Pada penelitian ini, metode klasifikasi yang digunakan yaitu C4.5, KNN, *Naive Bayes Classifier* (NBC) dan *Support Vector Machines* (SVM).

Algoritme C4.5 merupakan salah satu algoritme yang digunakan untuk membentuk pohon keputusan. Algoritme ini mempunyai kelebihan misalnya dapat mengolah data numerik dan diskret, dapat menangani nilai atribut yang hilang. Berikut persamaan untuk membentuk pohon keputusan dengan algoritme C4.5. Sebelum menghitung nilai *gain*, terlebih dahulu menghitung nilai *entropy*. Untuk menghitung nilai *entropy* menggunakan persamaan berikut:

$$Entropy(S) = \sum_{i=1}^{n} - pi * \log_2 pi \quad \text{.....(2)}$$

dimana, *S* adalah himpunan kasus, *n* adalah jumlah partisi *S*, *pi* proporsi dari *Si* terhadap *S*
Kemudian hitung nilai Gain dengan persamaan berikut :

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy(S_i) \quad \text{.....(3)}$$

dimana,
S = himpunan kasus
A = atribut
N = jumlah partisi *S*
|*Si*| = jumlah kasus pada partisi ke-i
|*S*| = jumlah kasus dalam *S*

KNN merupakan suatu pendekatan klasifikasi yang mencari semua data latih yang relatif mirip dengan data uji [8]. Prinsip kerja KNN yaitu mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga terdekatnya dalam data pelatihan. Berikut persamaan pencarian jarak menggunakan Euclidean [8] :

$$d(x,y) = \sqrt{\sum_{k=1}^{n}(x_k - x_k)^2} \quad \text{.....(4)}$$

Dimana matriks *d(x,y)* adalah jarak skalar dari kedua vektor *x* dan *y* dari matriks dengan ukuran *d* dimensi.
NBC merupakan sebuah metode klasifikasi yang menerapkan teorema *Bayes* yang berdasarkan nilai probabilitas [13]. NBC menyediakan algoritme yang efisien dan sederhana. Selain itu, NBC memiliki kelebihan yaitu hanya memeriksa training data sekali serta hanya memerlukan training data yang sedikit untuk mengestimasi parameter seperti rerata dan variansi variabel. NBC dirumuskan dengan persamaan berikut :

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \quad \text{.....(5)}$$

dimana,
X = data dengan kelas yang belum diketahui.
Y = hipotesis data *X* merupakan suatu kelas spesifik.
P(Y|X) = probabilitas hipotesis *Y* berdasar kondisi *X*.
P(Y) = probabilitas awal hipotesis *Y* (*prior probability*)
P(X|Y) = probabilitas *X* berdasar kondisi pada hipotesis *Y*.
P(X) = probabilitas dari *X*.

SVM merupakan algoritme klasifikasi yang menggunakan ruang hipotesis berupa fungsi-fungsi linear dalam sebuah ruang berdimensi fitur tinggi. SVM memiliki tujuan menemukan fungsi pemisah terbaik antara kelas. Fungsi pemisah paling baik apabila fungsi pemisah yang menghasilkan nilai margin paling besar antara dua vektor dari dua kelas yang berbeda dan berada ditengah-tengah kedua vektor tersebut. Dalam penelitian ini, fungsi pemisah yang dicari adalah fungsi linear. Fungsi linear dapat dicari dengan persamaan berikut :

$$g(x) = \varpi.x + w_0 \quad \text{.....(6)}$$

dimana,
ϖ = vektor untuk *hyperplane*
x = nilai vektor masukan
w_0 = bias yang merepresentasikan posisi relatif terhadapt pusat koordinat.

4 Evaluasi

Evaluasi dilakukan untuk mengukur kelayakan dari model yang dikembangkan, sehingga akan diketahui apakah layak atau tidak untuk diimplementasikan dalam klasifikasi SMS spam. Untuk mengukur kinerja dari model klasifikasi SMS spam yaitu dengan menghitung akurasi. Untuk menghitung akurasi dapat menggunakan *confusion matrix* sebagai pengukuran kinerja dari suatu model.

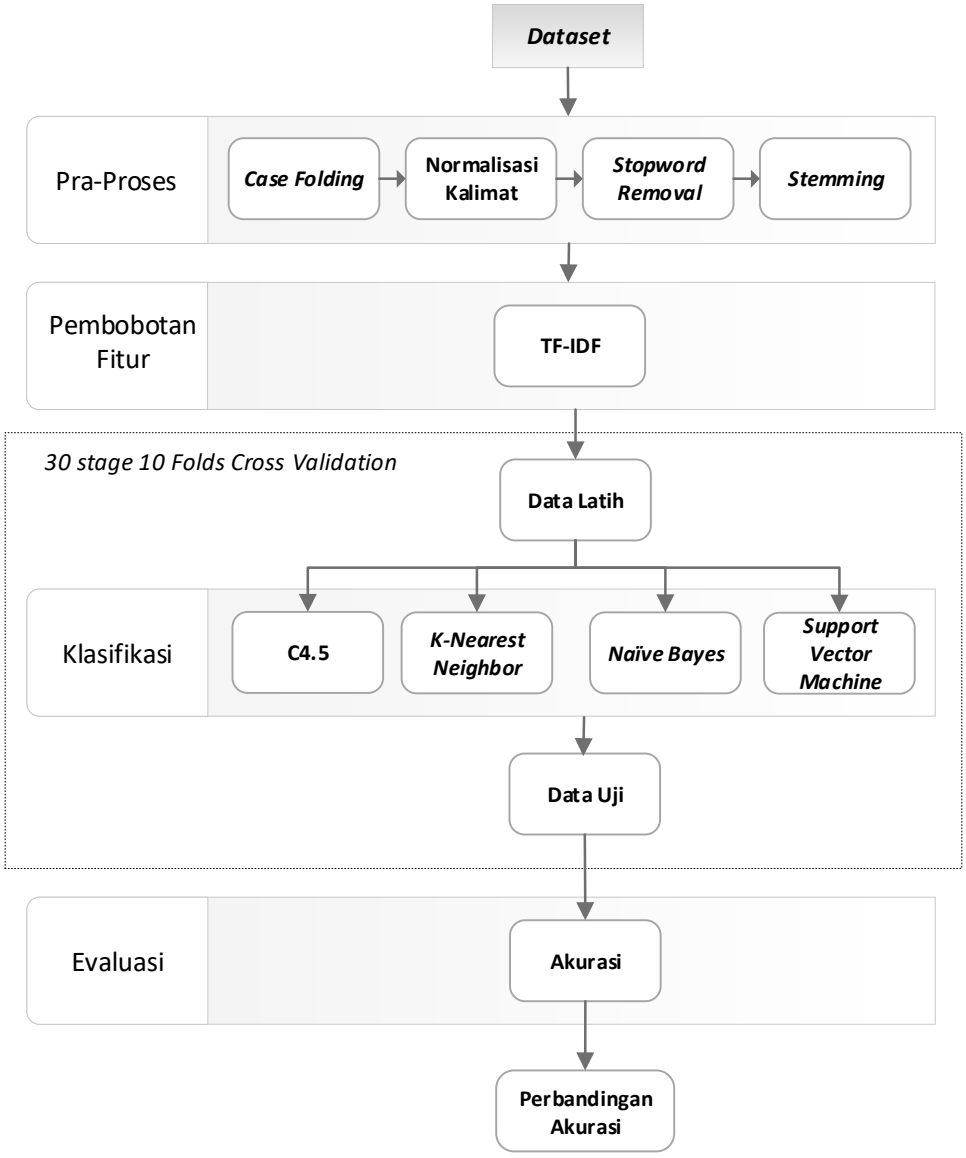
Tabel 1. *Confusion matrix* dengan dua kelas

Kelas	**Terklasifikasi**	
	Spam	*Ham*
Spam	True Positive (TP)	False Negative (FN)
Ham	False Positif (FP)	True Negative (TN)

Berdasarkan nilai *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN) pada tabel *Confusion matrix* maka dapat diperoleh nilai akurasi. Nilai akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar. Dengan kata lain, nilai akurasi merupakan perbandingan antara data yang terklasifikasi secara benar dengan keseluruhan data. Berikut persamaan untuk mencari nilai akurasi berdasarkan *confusion matrix*.

$$akurasi = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \quad \text{..... (7)}$$

Metode evaluasi yang digunakan yaitu *30-Stage 10-Cross Fold Validation*. Dimana metode *10-Cross Fold Validation* diulang sebanyak 30 kali (*30-Stage 10-Cross Fold Validation*) dengan pengacakan data latih dan data uji pada setiap iterasinya. Keseluruhan proses metode penelitian yang digunakan dapat dilihat pada Gambar 1 berikut ini.



Gambar 1. Diagram blok perbandingan algoritme klasifikasi

3. Pembahasan

Penelitian ini dilakukan menggunakan laptop dengan spesifikasi CPU *Intel Core* i5 2.60 GHz, RAM 8 GB dan dengan sistem operasi *Microsoft Windows 7 Profesional 64 bit*. Bahasa pemrograman yang digunakan yaitu python dengan bantuan pustaka sastrawi dan pustaka *sklearn*. Pustaka sastrawi berasal dari github yaitu https://github.com/sastrawi/sastrawi yang digunakan untuk mengimplementasikan proses *stemming*. Pustaka *sklearn* digunakan untuk mengimplementasikan proses pembobotan fitur, klasifikasi dan evaluasi.

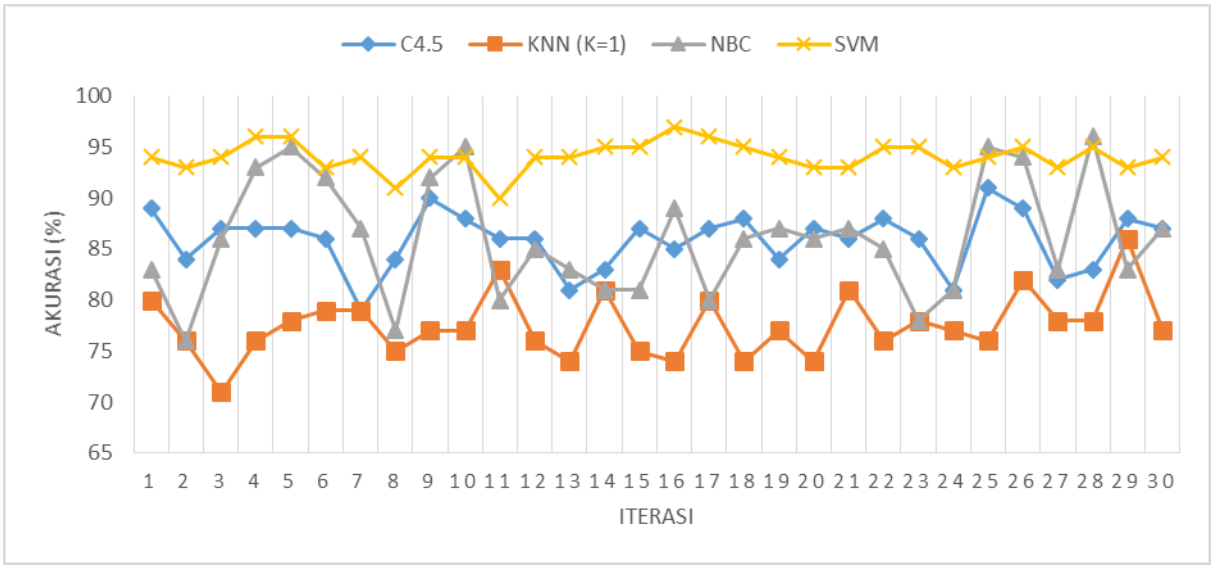
Setelah dilakukan klasifikasi terhadap *dataset* SMS spam dengan menggunakan algoritme C4.5, KNN, NBC dan SVM, maka hasil pengujian tingkat akurasi dengan pengujian *30 stage 10-fold cross validation* dapat dilihat pada Tabel 2.

Tabel 2. Perbandingan akurasi algoritme klasifikasi (%)

Stage	C4.5	KNN (K=1)	NBC	SVM
1	89,00	80,00	83,00	93,99
2	84,00	76,00	76,00	92,99
3	87,00	70,99	86,00	93,99
4	87,00	76,00	92,99	95,99
5	87,00	78,00	94,99	95,99
6	85,99	79,00	92,00	92,99
7	79,00	79,00	87,00	93,99
8	83,99	75,00	77,00	90,99
9	90,00	77,00	91,99	93,99
10	88,00	77,00	94,99	93,99
11	85,99	82,99	80,00	89,99
12	85,99	76,00	84,99	93,99
13	81,00	73,99	83,00	93,99
14	83,00	81,00	81,00	94,99
15	87,00	75,00	81,00	94,99
16	84,99	73,99	89,00	96,99
17	87,00	80,00	80,00	95,99
18	88,00	73,99	85,99	94,99
19	84,00	77,00	87,00	93,99
20	87,00	73,99	85,99	92,99
21	86,00	81,00	87,00	92,99
22	88,00	76,00	84,99	94,99
23	85,99	78,00	78,00	94,99
24	81,00	77,00	81,00	92,99
25	90,99	76,00	94,99	93,99
26	89,00	82,00	93,99	94,99
27	82,00	78,00	83,00	92,99
28	83,00	78,00	95,99	94,99
29	88,00	85,99	82,99	92,99
30	87,00	77,00	87,00	93,99
Mean	**85,86**	**77,50**	**86,10**	**94,06**

Pada Tabel 2 dapat dilihat bahwa algoritme SVM menghasilkan tingkat akurasi tertinggi dari pada algoritme C4.5, KNN dan NBC dari setiap pengujian sebanyak 30 kali (*stage*). Selain itu, hasil rerata nilai akurasi yang paling tinggi adalah algoritme SVM dengan rerata nilai akurasi rata-rata sebesar 94,06%. Ini menunjukkan bahwa tingkat akurasi algoritme SVM paling baik dari pada algoritme C4.5, KNN dan NBC dalam mengklasifikasikan SMS spam.

Untuk lebih memudahkan pengamatan perbedaan tingkat akurasi yang dihasilkan oleh keempat algoritme klasifikasi dengan pengujian *30 stage 10-fold cross validation* diatas, nilai akurasi yang dihasilkan ditampilkan dalam bentuk grafik yang dapat diliat pada Gambar 2.



Gambar 2. Grafik perbandingan akurasi algoritme klasifikasi

4. Simpulan

Hasil dari perbandingan algoritme klasifikasi antara C4.5, *K-Nearest Neighbor* (KNN), *Naïve Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM) didapatkan algoritme klasifikasi paling baik yaitu SVM dengan tingkat akurasi sebesar 94,06%. Kemudian, diikuti oleh algoritme NBC, C4.5 dan KNN dengan tingkat akurasi berturut-turut yaitu 86,10%, 85,86% dan 77,50%. Oleh karena itu, algoritme SVM sangat cocok untuk diimplementasikan pada klasifikasi SMS spam.

Daftar Pustaka

[1] N. Chaudhari, P. Jayvala, dan P. Vinitashah, *"Survey on Spam SMS filtering Using Data Mining Techniques,"* IJARCCE, vol. 5, no. 11, pp. 193–195, 2016.

[2] K. Mathew dan B. Issac, *"Intelligent spam classification for mobile text message,"* Proc. 2011 Int. Conf. Comput. Sci. Netw. Technol. ICCSNT 2011, vol. 1, pp. 101–105, 2011.

[3] J. Ma, Y. Zhang, J. Liu, K. Yu, dan X. Wang, *"Intelligent SMS Spam Filtering Using Topic Model,"* Intell. Netw. Collab. Syst. (INCoS), 2016 Int. Conf. IEEE, pp. 380–383, 2016.

[4] I. N. Dewi dan C. Supriyanto, "Klasifikasi Teks Pesan Spam Menggunakan Algoritma Naïve Bayes," Simantik 2013, vol. 2013, no. November, pp. 156–160, 2013.

[5] B. Indiarto, "Klasifikasi Sms Spam Dengan Metode *Naive Bayes Classifier* Untuk Menyaring Pesan Melalui Selular," J. Telemat. *MKOM*, vol. 8, no. 2, pp. 167–172, 2016.

[6] T. H. Apandi dan C. A. Sugianto, "Penyaringan *Spam Short Message Service* Menggunakan Support Vector Machine," Semin. Nas. Teknol. Inf. dan Komun. Terap. 2015, pp. 111–116, 2015.

[7] S. N. D. Pratiwi dan B. S. S. Ulama, "Klasifikasi Email Spam dengan Menggunakan Metode *Support Vector Machine* dan *k-Nearest Neighbor*," J. SAINS dan SENI ITS, vol. 5, no. 2, pp. 344–349, 2016.

[8] W. Gata dan Purnomo, "Akurasi *Text Mining* Menggunakan *Algoritma K-Nearest Neighbour* pada *Data Content SMS-Gateway*," J. Format, vol. 6, no. 5, pp. 1–5, 2017.

[9] S. Youn dan D. McLeod, "*A comparative study for email classification,"* Adv. Innov. Syst. Comput. Sci. Softw. Eng., pp. 387–391, 2007.

[10] Z. R. Mair dan A. Ashari, "Aplikasi untuk Identifikasi *Short Message Service (SMS) Spam* Berbasis Android*," BIMIPA, vol.* 24, no. 3, Jun. 2014.

[11] M. Y. Nur dan D. D. Santika, "Analisis Sentimen pada Dokumen berbahasa Indonesia dengan pendekatan Support Vector Machine," *Konf. Nas. Sist. dan Inform.*, pp. 9–14, 2011.

[12] P. Widodo, J. A. Putra, S. Afiadi, A. Z. Arifin, dan D. Herumurti, "Klasifikasi Kategori Dokumen Berita Berbahasa Indonesia dengan Metode Kategorisasi Multi-Label Berbasis *Domain Specific Ontology*," J. Teknosains, vol. II, no. 2, pp. 101–112, Aug. 2017.

[13] S. Raschka, *"Naive Bayes and Text Classification I - Introduction and Theory," CoRR*, vol. abs/1410.5, Oct. 2014.