

# Analisis Klasifikasi SMS Spam Menggunakan Logistic Regression

**Ferin Reviantika**

Jurusan Teknik Informatika  
Universitas Muhammadiyah Malang  
Malang, Indonesia  
ferinurin27@gmail.com

**Yufis Azhar**

Jurusan Teknik Informatika  
Universitas Muhammadiyah Malang  
Malang, Indonesia  
yufis@umm.ac.id

**Gita Indah Marthasari**

Jurusan Teknik Informatika  
Universitas Muhammadiyah Malang  
Malang, Indonesia  
gita@umm.ac.id

*Abstract- SMS or Short Message Service is usually found on cell phones. SMS is divided into two categories, namely SMS spam and SMS non-spam (ham). Spam SMS is an SMS that is annoying to phone users because it tends to contain messages that are not important such as promos and scams. Meanwhile, non-spam SMS (ham) tend to contain important SMS, such as messages from previous users. In this study, the classification of spam SMS and non-spam SMS (ham) was carried out using the logistic regression method. The purpose of this study is to distinguish or classify between spam and non-spam SMS (ham). The dataset in this study amounted to 1143 data, there are two columns, namely the text column and the label column. The number for spam messages is 566 messages and the number for non-spam messages is 577. The proposed method gets a better accuracy of 95%.*

*Keywords— SMS, SMS spam, SMS non spam (ham), Logistic regression, Classification.*

## I. PENDAHULUAN

Short Message Service atau disingkat dengan SMS merupakan salah satu bagian dari layanan pesan perusahaan komunikasi yang ada pada telepon seluler atau handphone [1]. Rata – rata SMS hanya dapat mengirimkan pesan berupa teks dan hanya dapat digunakan untuk panggilan telepon. Fasilitas SMS pada telepon seluler tidak hanya dapat bertukar pesan dengan orang terdekat saja, namun juga mengirimkan pesan kepada orang yang belum mengenal satu sama lain, seperti menawarkan produk, promosi, jasa dan masih banyak lagi. SMS juga banyak digunakan untuk kasus penipuan.

SMS merupakan salah satu sistem yang memiliki jumlah pengguna terbesar di dunia. Di Indonesia sendiri pengguna SMS juga tergolong masih tinggi. Dikutip dari <https://www.cnbcindonesia.com> bahwa Indonesia punya 37 juta pelanggan yang teregistrasi dan 25 – 28 juta pengguna aktif. Tetapi SMS masih memiliki kelemahan dalam keamanan sistem. SMS dibagi menjadi dua kategori yaitu SMS spam dan SMS non spam atau ham. SMS spam adalah pesan – pesan yang tidak bermanfaat, sedangkan SMS non spam atau ham adalah pesan – pesan asli atau bersifat penting [1]. Hal ini memicu kekhawatiran dalam ruang lingkup pribadi karena biasanya SMS digunakan untuk pertukar pesan secara pribadi atau rahasia. Semakin banyaknya SMS spam yang tidak berguna sehingga terkadang mengganggu dan membuat ketidaknyamanan dalam penggunaan telepon seluler.

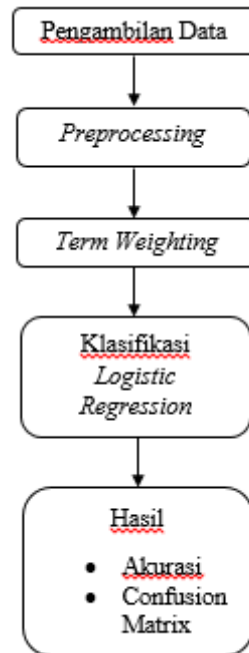
Klasifikasi adalah proses mencari model dengan cara mengkategorisasikan sehingga dapat memprediksi kelas yang tidak berlabel [2]. Penelitian tentang klasifikasi SMS sudah banyak dilakukan seperti penelitian terdahulu oleh Ika Novita dkk tentang klasifikasi teks spam menggunakan metode naïve bayes diperoleh hasil berupa akurasi sebesar 84,40%, recall 88,09% dan precision 45,76% [3]. Penelitian oleh Basuki dkk tentang klasifikasi teks *tweets* tindak kejahatan menggunakan berbahasa Indonesia dengan metode naïve bayes diperoleh hasil klasifikasi menggunakan metode *lexical* atau *bag of words* sebesar 79,25% sedangkan pada fitur sintaktik adalah sebesar 88,1398% [4]. Penelitian oleh Sravya dkk yang berjudul *mobile sms spam filter techniques using machine learning techniques* dengan dilakukan menggunakan banyak metode salah satunya menggunakan *Logistic Regression*, akurasi yang didapatkan sangat tinggi yaitu 97,8% [5].

Berdasarkan uraian diatas dapat dilihat metode terdahulu yaitu *naïve bayes* masih memiliki hasil akurasi yang kurang tinggi dalam melakukan klasifikasi antara SMS spam dan non spam (ham). Dan metode *logistic regression* memiliki hasil yang lebih baik. Hal tersebut yang memicu dilakukannya

penelitian ini untuk mempelajari klasifikasi SMS spam menggunakan metode yang berbeda, yaitu Logistic Regression [6].

## II. METODE PENELITIAN

Metode penelitian adalah alur penelitian yang dilakukan pada penelitian menggunakan *logistic regression* untuk klasifikasi SMS spam, dimana hal ini meliputi beberapa tahapan dimulai dari *literature view*, pengambilan data, *preprocessing*, TF – IDF, analisis hasil, dan uji coba dataset. Berikut adalah gambaran penelitian :



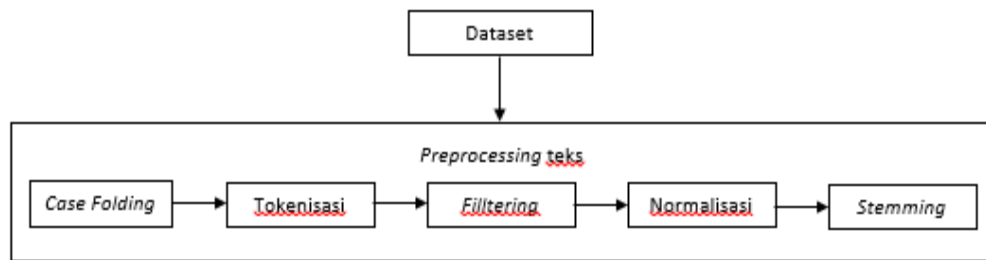
Gambar 1. Flowchart dan tahap penelitian

### A. Pengambilan Data

Pada penelitian ini data diperoleh dari <https://github.com> yang berjudul klasifikasi spam SMS milik Kuncahyo Setyo Nugroho. Data ini di upload pada tanggal 15 November 2019 sebanyak 1143 *record*. Data terdiri dari dua atribut yaitu atribut nomor, atribut label yang dibedakan menjadi tiga kelas yaitu normal diindikasikan dengan nilai 0, penipuan diindikasikan dengan nilai 1, promo diindikasikan dengan nilai 2 dan atribut *message* atau contoh pesan SMS. Karena penelitian ini bertujuan untuk klasifikasi SMS spam dan non spam (ham), sehingga menyatukan kategori dari pesan spam dimana dalam dataset dijelaskan pada kolom label bernilai 1 dan 2 digabungkan menjadi berlabel 1 yaitu spam.

### B. Preprocessing

Preprocessing merupakan proses awal dalam pengolahan data masukan sehingga akan menghasilkan data dengan format yang sesuai dan siap diproses pada tahap selanjutnya [7]. Tujuan dilakukannya *preprocessing* agar data yang digunakan lebih presisi dan memudahkan dalam menjalankan klasifikasi. Tahapan *preprocessing* dapat dilihat dalam gambar 3.



Gambar 2. Tahapan *Preprocessing*

### 1) *Case Folding*

Pada tahap penelitian ini akan mengubah atau mengkonversi *text* menjadi bentuk standart. Sehingga data masukan atau data primer akan diubah menjadi *lowercase* atau huruf kecil [8].

### 2) *Tokenisasi*

Untuk memudahkan *preprocess text* dalam tahap tokenisasi memisahkan teks menjadi token token [9]. Sehingga perlu menghapus tanda baca yang tidak perlu, menghapus semua *extra space*, menghapus semua kata yang berjumlah 1, *line breaks*, dan *tabs*.

### 3) *Filtering*

Dalam proses *Filtering* dilakukan proses *stopwords*, proses *stopword* akan menghapus kata-kata yang tidak memiliki arti yang signifikan atau biasa disebut *meaning less* seperti “ny”, “gk”, “dg”, “sih”, “hehe” dll [10].

### 4) *Normalisasi*

Pada tahap ini akan dilakukan penyeragaman kata yang memiliki makna yang sama, dapat diakibatkan dari penulisan yang salah, penyingkatan kata, ataupun bahasa gaul [11].

### 5) *Stemming*

Proses *stemmer* yaitu mengembalikan kata yang berimbuhan menjadi kata induk [12].

## C. *Term Weighting*

*Term weighting* pada penelitian ini adalah menggunakan nilai TF dan IDF. TF (*Term frequency*) adalah metode sederhana dalam pembobotan setiap *term* [13].

$$tf_{ij} = \frac{f_d(i)}{N f_d(j)} \quad (1)$$

Sedangkan IDF adalah perhitungan *term* yang di distribusikan dalam sebuah dokumen [14].

$$idf(t, d) = \log \left( \frac{N}{df(t)} \right) \quad (2)$$

## D. *Klasifikasi Logistic Regression*

*Logistic Regression* adalah bagian dari metode data mining yang penggunaannya untuk menganalisis data yang mendeskripsikan antara satu variabel respon (dependen) atau lebih variabel *predictor* [15]. Metode ini digunakan saat variabel *predictor* (y) memiliki skala kategorik atau nominal yang terdiri dari dua (biner) atau lebih kategori. Sehingga metode ini dibuat untuk memastikan bahwa, apa pun perkiraan yang terjadi, selalu berada di antara 0 dan 1 [16].

Model nilai *w* dapat dilihat pada persamaan (3)

$$W = b + w_1x_1 + w_2x_2 + w_3x_3 + \dots + w_nx_n \quad (3)$$

Nilai *w* yang diperoleh kemudian dipetakan menggunakan fungsi *logistic regression*.

$$P = \frac{1}{1+e^{-(w)}} \quad (4)$$

Persamaan (4) merupakan persamaan umum yang dapat diubah menjadi

$$P = \frac{1}{1+e^{-(b+w_1x_1+w_2x_2+w_3x_3+\dots+w_nx_n)}} \quad (5)$$



Dari gambar 5 dan gambar 6, metode *logistic regression* menghasilkan nilai TP atau pesan yang benar diprediksi sebagai pesan Ham sebanyak 110, FP atau pesan yang sebenarnya bernilai Ham namun diprediksi Spam sebanyak 6, FN atau pesna yang bernilai Spam namun diprediksi Ham sebanyak 5, dan TN atau pesan yang benar diprediksi sebagai pesan Spam sebanyak 108. Sehingga dapat dapat diketahui nilai *precision*, *recall* dan *f1-score*.

	Predicted Ham	Predicted Spam
Actual Ham	110	6
Actual Spam	5	108

Gambar 5. Confusion matrix

	precision	recall	f1-score
0	0.96	0.95	0.95
1	0.95	0.96	0.95
accuracy			0.95

Gambar 6. Classification report

Akurasi adalah nilai dari perbandingan kasus yang mengidentifikasi benar dari jumlah kelas [18]. Sedangkan *precision* dan *recall* mengukur ketepatan dan kelengkapan model klasifikasi dalam penelitian ini [19]. Yang terakhir adalah *f1-score* adalah perbandingan nilai rata-rata dari nilai *precision* dan nilai *recall* [18]. Pada gambar 5 dapat dilihat hasil dari *precision*, *recall* dan *f1-score*. Nilai 0 pada gambar 5 menandakan hasil untuk pesan non spam (ham) dan nilai 1 menandakan hasil untuk pesan spam. Hasil dari *Classification report logistic regression* juga dapat dilihat pada tabel 1.

Penelitian ini juga membandingkan metode yang diusulkan dengan metode lain seperti *naïve bayes*. Dari tabel 1 dapat dilihat metode *logistic regression* memiliki hasil yang lebih unggul dibandingkan metode *naïve bayes* diukur dari *precision*, *recall*, *f1-score*, dan *Accuracy*.

Tabel 1.  
Hasil Perbandingan Metode

Metode	Precision (%)		Recall (%)		F1-score (%)		Accuracy (%)
	0	1	0	1	0	1	
LR	0.96	0.95	0.95	0.9	0.95	0.95	0.95
NB	0.95	0.91	0.9	0.9	0.93	0.93	0.93

#### IV. KESIMPULAN

Berdasarkan hasil penelitian klasifikasi SMS spam dan Non spam dengan menggunakan metode *logistic regression*, didapatkan hasil yang baik dan tingkat akurasi yang tinggi pada metode *logistic regression* pada proses pengujian klasifikasi. Dengan menggunakan perbandingan data *train* dan *test* sebanyak 80:20 akurasi yang didapat lebih baik dari metode pembanding yaitu 95%.

## REFERENSI

- [1] B. Indarto, "KLASIFIKASI SMS SPAM DENGAN METODE NAIVE BAYES CLASSIFIER UNTUK MENYARING PESAN MELALUI SELULAR," vol. 8, no. 2, pp. 167–172, 2016.
- [2] A. Setiawan, I. F. Astuti, and A. H. Kridalaksana, "KLASIFIKASI DAN PENCARIAN BUKU REFERENSI AKADEMIK MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER ( NBC ) ( STUDI KASUS : PERPUSTAKAAN DAERAH PROVINSI KALIMANTAN TIMUR )," vol. 10, no. 1, 2015.
- [3] I. N. Dewi, C. Supriyanto, F. I. Komputer, and U. D. Nuswantoro, "Klasifikasi Teks Pesan Spam Menggunakan Algoritma Naive Bayes," vol. 2013, no. November, pp. 156–160, 2013.
- [4] S. Basuki, S. Maghfiroh, and Y. Azhar, "Klasifikasi Tweets Tindak Kejahatan Berbahasa Indonesia Menggunakan Naive Bayes," vol. 2, no. 7, pp. 933–944, 2020.
- [5] G. S. Sravya and G. Pradeepini, "Mobile Sms Spam Filter Techniques Using Machine Learning Techniques," vol. 9, no. 03, pp. 384–389, 2020.
- [6] I. T. Utami, "Analisis Regresi Logistik Biner Untuk Mengklasifikasi Penderita Hipertensi Berdasarkan Kebiasaan Merokok Di RSUD Mokopido Toli-Toli Binary Logistic Regression Analysis For Classification Of Hypertension Patients Based On Smoking Habits In Mokopido Toli-Toli Hospital," vol. 7, no. 3, pp. 341–348, 2018.
- [7] D. B. Setyohadi and F. A. Kristiawan, "Perbaikan performansi klasifikasi dengan preprocessing iterative partitioning filter algorithm 1," vol. 14, no. 01, pp. 12–20, 2017.
- [8] F. Ratnawati, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," 2018.
- [9] A. H. Yunas and M. Fikry, "Klasifikasi Tweet E-Commerce dengan Menggunakan Metode Support Vector Machine," vol. 4, no. 2, pp. 50–55, 2018.
- [10] R. Melita *et al.*, "( TF-IDF ) DAN COSINE SIMILARITY PADA SISTEM TEMU KEMBALI INFORMASI UNTUK MENGETAHUI SYARAH HADITS BERBASIS WEB ( STUDI KASUS : SYARAH UMDATIL AHKAM )," vol. 11, no. 2, 2018.
- [11] K. S. Setyawati *et al.*, "Aplikasi Sentiment Analysis Terhadap Pelaksanaan Pembelajaran Jarak Jauh Universitas Kristen Petra Dengan Metode Naive Bayes Classifier."
- [12] A. Y. Permana, "IMPLEMENTASI STEMMING PORTER KBBI UNTUK KLASIFIKASI TOPIK SOAL UJIAN NASIONAL BAHASA INDONESIA MENGGUNAKAN ALGORITMA NAIVE BAYES," 2017.
- [13] 2016 Kurniawati, "Term weighting berbasis indeks kelas menggunakan metode tf.idf.ics," 2016.
- [14] T. Trisna *et al.*, "ANALYSIS AND DETECTION OF HOAX CONTENTS IN INDONESIAN NEWS BASED ON MACHINE LEARNING," vol. 4, no. 1, 2019.
- [15] C. Science, A. Bimantara, and T. A. Dina, "Klasifikasi Web Berbahaya Menggunakan Metode Logistic Regression," vol. 4, no. 1, pp. 978–979, 2018.
- [16] O. S. Balogun, T. J. Akingbade, and A. Akinrefon, "Evaluation Of Logistic Regression In Classification Of Drug Data In Kwara State," no. May 2014, 2013.
- [17] H. Saiyar, "Aplikasi Diagnosa Penyakit Tuberculosis Menggunakan Algoritma Naive Bayes," vol. 5, no. 5, pp. 498–502, 2018.
- [18] M. Athaillah, Y. Azhar, and Y. Munarko, "Perbandingan Metode Klasifikasi Berita Hoaks Berbahasa Indonesia Berbasis Pembelajaran Mesin," vol. 2, no. 5, pp. 675–682, 2020.
- [19] D. H. Kalokasari, I. M. Shofi, and A. H. Setyaningrum, "IMPLEMENTASI ALGORITMA MULTINOMIAL NAIVE BAYES CLASSIFIER PADA SISTEM KLASIFIKASI SURAT KELUAR ( Studi Kasus : DISKOMINFO Kabupaten Tangerang )," vol. 10, no. 2, 2017.