

## KLASIFIKASI KOMENTAR SPAM PADA INSTAGRAM MENGUNAKAN METODE *SUPPORT VECTOR MACHINE*

Agus Tiyansyah Syam<sup>1</sup>, Amin Fahri<sup>2</sup>, Bayu Saputra<sup>3</sup>, Rian Ferdian Maulana<sup>4</sup>, Selvi Dwiani<sup>5</sup>, Wildan Gunawan Holid<sup>6</sup>, Ricky Firmansyah, ST., M.Kom<sup>7</sup>

<sup>1,2,3,4,5,6,7</sup> *Fakultas Teknik Informatika, Universitas ARS*

*Jl. Sekolah Internasional No. 1 - 2, Antapani, Bandung, Jawa Barat, 40282.*

*Email : \*<sup>1</sup>agustiyaninfo@gmail.com, \*<sup>2</sup>aminfahri512@gmail.com, \*<sup>3</sup>yukibayu7@gmail.com, \*<sup>4</sup>rianferdian2@gmail.com, \*<sup>5</sup>piewdwian@gmail.com, \*<sup>6</sup>wildangunawanholid29@gmail.com \*<sup>7</sup>rickyfirmanars@gmail.com*

### **Abstrak**

Para publik figur, terutama aktor dan artis menggunakan Instagram (IG) sebagai salah satu media sosial berbasis foto atau video untuk mempromosikan kegiatan kepada penggemarnya. Para penggemar dapat memfollow, melihat postingan-postingan dan komentar para idolanya. Permasalahannya banyak sekali komentar spam pada postingannya baik berupa hal negatif maupun komentar jualan bahkan link ke suatu website berbahaya tertentu dan lain sebagainya. Untuk mengetahui bahwa suatu komentar adalah spam dapat dilakukan dengan melihat satu persatu dari komentar yang ada, namun hal ini akan membutuhkan tenaga dan memakan waktu yang banya (Cristianto, Antonius Rahmat, dan Yuan Lukito, 2017). Maka dari itu diperlukan sebuah sistem yang dapat mengklasifikasi bahwa suatu komentar adalah spam atau nonspam secara otomatis. Klasifikasi adalah sebuah proses dalam menentukan model atau fungsi yang mampu menjelaskan atau membedakan konsep (kelas data). Tujuannya adalah untuk memprediksi kelas yang tidak diketahui dalam suatu objek. Metode yang digunakan adalah Support Vector Machine (SVM). Data komentar yang digunakan pada penelitian ini dikumpulkan dari komentar-komentar pada foto atau video yang dibagikan oleh aktor atau artis yang memiliki pengikut (followers) diatas 1 juta pengikut.

**Kata Kunci:** *Klasifikasi Spam; Komentar Spam; Instagram; Support Vector Machine (SVM); Media Sosial*

### **Abstract**

*Public figures, especially actors and artists use instagram (ig) as one of those social-based photographs or videos to promote activities to fans. Fans can follow, seeing posts and comments by their idols. The thing is, there are so many spamming comments on his posts of both negative and malicious comments that selling even links to a certain dangerous website and so on. To know that a comment is spam can be done by seeing one at a time, but it will require energy and a great amount of time (cristianto, antonius grace, and yuan lukito, 2017). Therefore a system that can classify a comment as being spam or nonspam automatically. A classification is a process in determining a model or function that is able to explain or discern concepts (data class). The goal is to predict an unknown class in an object. A method of using is the support for vector machine (SVM). The commentaries used in this study are collected from comments on photos or videos Shared by actors or artists who have followers (followers) over 1 million followers.*

**Keywords:** *Spam Classification; Spam Comments; Instagram; Support Vector Machine (SVM); Social Media*

## **1. PENDAHULUAN**

Di era berkembangnya penggunaan Internet, penggunaan Internet memberikan dampak langsung terhadap penggunaan media sosial. Pada tahun 2012, pengguna Internet di Indonesia

mencapai 63 juta orang, 95 persen di antaranya menggunakan Internet untuk mengakses media sosial seperti Facebook, Twitter, Instagram, dan lain-lain.

Pesatnya perkembangan dunia teknologi dan informasi dalam beberapa tahun terakhir, menjadikan internet sebagai alat komunikasi yang banyak diminati oleh masyarakat. Hal inilah yang melatarbelakangi perubahan komunikasi konvensional menjadi modern dan serba digital. Perkembangan ini pun menjadi semakin pesat setelah internet mulai dapat diakses melalui telepon seluler dan muncul istilah telepon cerdas (*smartphone*). *Smartphone* memberikan fasilitas yang beraneka ragam, mulai dari SMS, MMS, *email*, *chatting*, *browsing*, serta fasilitas sosial media. Berdasarkan data pengguna internet Indonesia, sampai dengan akhir tahun 2005 pengguna internet Indonesia mencapai 16 juta pengguna, naik hampir 50% dibandingkan dengan tahun 2004 yang hanya 11 juta pengguna. Sementara pada akhir tahun 2011 pengguna internet Indonesia telah mencapai lebih dari 50 juta pengguna.

Instagram (IG) merupakan aplikasi berbasis android, iOS, maupun web yang memiliki banyak pengguna di penjuru dunia. Dimana cara kerja dari aplikasi IG ini adalah pengguna dapat *men-posting* foto atau video ke aplikasi IG sehingga postingan tersebut dapat dilihat oleh orang lain.

Di Indonesia sendiri aplikasi IG cukup banyak mempunyai peminat dari berbagai kalangan, juga berbagai umur. Data dari Napoleon Cat, mencatat pada bulan Januari hingga Mei 2020 pengguna IG di Indonesia telah mencapai 69,2 juta.

Tak beda halnya dengan para tokoh, artis, atau influencer di Indonesia mereka juga menggunakan IG sebagai media sosial. Beberapa diantaranya bahkan memiliki followers 5 juta bahkan lebih. Selain sebagai alat untuk bersosial secara daring IG pun dimanfaatkan sebagai media mencari keuntungan, seperti *endorse*, dan berjualan. Namun semakin besarnya pengguna aplikasi IG tak semua hal bisa dikontrol secara penuh, misalnya komentar. Sering kali dijumpai komentar yang terindikasi spam pada postingan akun - akun dengan jumlah *followers* yang besar. Spam komentar tersebut bermacam - macam dari mulai karakter yang tak jelas sampai komentar jualan yang mana hal tersebut tidak berkaitan dengan postingan yang ada.

Pada penelitian ini pokok permasalahan yang dibahas adalah mengenai bagaimana cara mengidentifikasi komentar yang terindikasi spam atau tidak dengan menggunakan algoritman *Support Vector Machine* (SVM).

## 2. METODE PENELITIAN

*Support vector machine* adalah suatu teknik untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. *Support vector machine* berada dalam satu kelas dengan *Artificial Neural Network* dalam hal fungsi dan kondisi permasalahan yang bisa diselesaikan. Keduanya masuk dalam kelas *supervised learning* (Santoso 2007).

*Support vector machine* adalah metode yang menganalisis data dan mengenali pola yang digunakan untuk klasifikasi data. Sistem kerja *Support vector machine* dengan cara memasukan data tertentu kedalam sebuah sistem kemudian memprediksikannya. Dari data tersebut akan di kelompokkan kedalam dua kelas yang berbeda. Data baru tersebut akan diklasifikasikan kedalam kelas yang satu atau kelas yang lain. Dari hal tersebut maka akan jelas bahwa data tersebut akan dikelompokkan ke dalam kelas yang sesuai (Lidya, Sitompul, & Efendi, 2015).

*Support vector machine* adalah metode terbaik yang bisa dipakai dalam permasalahan klasifikasi. Konsep *Support vector machine* bermula dari masalah klasifikasi dua kelas sehingga membutuhkan training set positif dan negatif. *Support vector machine* berusaha menemukan pemisah terbaik untuk memisahkan ke dalam dua kelas dan memaksimalkan margin antara dua kelas. Pada banyak kasus, data tidak bisa diklasifikasi menggunakan metode linier *Support vector machine*, sehingga dikembangkanlah fungsi kernel untuk mengklasifikasikan data dalam bentuk non-linier (Pratama, Wihandika, & Ratnawati, 2018).

### 2.1 Sumber Data

Sumber data yang digunakan pada penelitian ini adalah kumpulan komentar yang dipilih dari 5 (lima) akun IG yang memiliki followers lebih dari 1 juta pengguna dan sudah verified. Adapun ke-lima akun tersebut yaitu @fadiljaidi, @dagelan, @lambe\_turah, @pstOre, dan @raffinagita1717.

Dari ke-lima akun tersebut dipilih per-akun satu postingan dan perpostingan diambil 100 komentar. Sehingga data terkumpul sejumlah 500 data.

### 2.2. Langkah Analisis

Langkah analisis yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Tahap pemrosesan data (Data cleaning)
2. Tahap *preprocessing*

- Information gain yaitu mengukur efektifitas suatu atribut dalam mengklasifikasi data.
3. Mendapatkan komentar yang sudah di-*preprocessing*.
  4. Menganalisis karakteristik dari komentar spam dan ham.
  5. Klasifikasi komentar menggunakan metode SVM dengan SVM linier.

### 3. HASIL DAN PEMBAHASAN

Pada bagian ini, dijelaskan hasil penelitian dengan tahap yang tersebut pada metode penelitian.

#### 3.1. Tahap Pemrosesan data

Pada langkah ini dilakukan cleaning data dengan melakukan penghapusan pada karakter-karakter khusus, angka, dan data kosong. Setelah dilakukan data cleaning dihasilkan tetap 500 data yang terkumpul. Data dapat dilihat pada Tabel berikut:

Tabel 1. Skala Penilaian Survey

No	Username	Nama Kelas dan Jumlah
1.	@fadiljaidi	Spam 273 Bukan Spam 4188
2.	@dagelan	Spam 80, Bukan Spam 383
3.	@lambe_turah	Spam 451 Bukan Spam 1391
4.	@pst0re	Spam 475 Bukan Spam 2360
5.	@raffinagita1717	Spam 372 Bukan Spam 1438
Jumlah Spam 1.651		Jumlah Bukan Spam 9.760
Total Keseluruhan 11.411		

#### 3.2. Text Preprocessing

Tahapan ini adalah tahap *screening* data yang bertujuan mendapatkan karakteristik komentar dengan melihat kata yang muncul pada komentar dan frekuensi kemunculan kata dan untuk meningkatkan ketepatan klasifikasi dari data. Proses *case folding*, *stemming*, *stopwords* dan *tokenizing* tidak dilakukan. Hal ini dikarenakan kata penghubung, kata yang bukan kata dasar, dan kata yang tidak relevan bisa menjadi variabel yang signifikan dalam penelitian pada penelitian ini. Sehingga tahapan yang dilakukan hanyalah *information gain*, yang meliputi pembuatan *wordlist*. Dengan menggunakan bantuan python, sublime text dan R didapat hasil seperti pada Tabel berikut:

Tabel 2. Kumpulan Macam Kata dalam *Wordlist*

No	Kata
10	Jit
24	uwuwuw
64	Ginjal
....	....
....	....
189	Anu

Dari Tabel 2. Ditemukan terdapat 769 macam kata yang muncul pada keseluruhan *dataset*. Tahap selanjutnya yaitu pemilihan 100 jumlah kata yang didasarkan pada pembobotan *feature selections*. Untuk menyelesaikan tahapan *feature selections*, digunakan software Sublime Text (Unregistered Vesion) dengan menambahkan formula pembobot yang memenuhi persamaan berikut:

$$D = |\Pr(w|S) - \Pr(w|H)|$$

Tahapan ini juga mencakup perhitungan *selection of word* dan hasil *feature vectors content* dapat dilihat pada Tabel berikut:

Tabel 3. Hasil *Feature Vectors Content* Dengan Bobot *Spamcity*

No	Kata	Pr(w H)	Pr(w S)	D	Spamcity
1	Bisa	0.34	0	0.34	0
2	Gabung	0.24456	0.2113	0.4910	0.49976
3	Lihat	0.52	0.356	0.328	
...	...	...	...	...	...
...	...	...	...	...	...
48	lanjut	0.2	0.283	0.077	0.126
49	beban	0.07346	0.0543	0.0854	0.343
50	beli	0.034	0.2188	0.0564	0.034

Dari tabel 3. Dapat disimpulkan bahwa  $\Pr(w|H)$  merupakan nilai probabilitas kemunculan kata tersebut pada komentar ham. Sedangkan  $\Pr(w|S)$  merupakan nilai probabilitas kemunculan kata tersebut pada komentar spam. D adalah nilai yang mutlak dari perbedaan probabilitas komentar spam ( $\Pr(w|S)$ ) dan ham ( $\Pr(w|H)$ ). Serta *spamcity* adalah nilai kecenderungan kata pada data yang digunakan untuk menentukan kelas dari kata tersebut. Jika nilai *spamcity* dekat dengan nol maka kata tersebut merupakan indikator ham yang baik, sedangkan jika semakin mendekati satu maka merupakan indikator spam yang baik.

Tabel 4. Hasil *Word Frequencies*

Kata	Frekuensi
Mulai	459
Murah	732
...	...
...	...
Langsung	548
kak	865

Untuk melihat kata-kata yang dapat menyusun salah satu kelas, dapat dilakukan dengan melihat nilai nilai bobot  $Pr(w|H)$  dan  $Pr(w|S)$ . Jika nilai bobot, baik  $Pr(w|H)$  dan  $Pr(w|S)$ , lebih dari nol tetapi kurang dari satu, maka kata tersebut muncul di kedua komentar. Jika nilai bobot  $Pr(w|H) = 0$  maka kata tersebut hanya ada pada komentar spam dan sebaliknya. Sedangkan frekuensi dari 100 kata yang menjadi variabel ditampilkan pada Tabel 4.

### 3.3 SVM Linier

Parameter yang digunakan pada SVM linier adalah parameter C yang mana nilai parameter C yang akan memberikan hasil terbaik adalah  $C = 1$  nilai ini didasarkan pada metode trial error dengan mencoba nilai C dari 1 hingga 10. Hasil metode trial error ditunjukkan pada Tabel 5.

Tabel 5. Kecepatan Parameter C dengan *Trial Error*

C	Ketepatan
1	97.33%
2	97.48%
3	97.50%
4	97.68%
5	97.50%
6	97.90%
7	97.56%
8	97.73%
9	97.50%
10	97.58%

Berdasarkan hasil pada Tabel 5. Dapat diketahui bahwa ketepatan tertinggi diberikan oleh  $C = 1$  yaitu sebesar 97.33%. Sehingga analisis akan dilakukan dengan menggunakan  $C = 1$  agar mendapat hasil yang optimal. Berikut adalah hasil analisis dengan menggunakan SVM linier dengan  $C = 1$ .

Tabel 6. Hasil Kecepatan Klasifikasi SVM-Kernel Linier Pada Semua Fold

Kecepatan Klasifikasi	Precision	Recall	Error
97.33%	97.33%	97.33%	2.8%

Dengan data pada Tabel 6. didapatkan data bahwa ketepatan klasifikasi yang diberikan oleh SVM dengan SVM Linier adalah sebesar 97.33%. Nilai precision adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan hasil yang diberikan oleh sistem, dari hasil Tabel 6. diketahui bahwa hasil precision terbaik adalah 97.33% hal itu berarti sistem mampu mengklasifikasikan data dengan ketepatan hingga 97.33%.

Sedangkan untuk nilai recall adalah ketepatan sistem untuk mengklasifikasi data dengan benar, dari hasil Tabel 6. diketahui bahwa hasil recall terbaik adalah 97.33% artinya adalah sistem mampu mengklasifikasikan data dengan benar dengan ketepatan 97.33%. Error maksimum yang diberikan sebesar 2.8% artinya dari 500 komentar yang salah di klasifikasikan ada sebesar 14 komentar, dengan rincian seperti pada Tabel 7.

Tabel 7. Hasil *Confusion Matrix* SVM-Kernel Linier Pada Semua Fold

Confusion Matrix		Ham	Spam	Total
	Pengamatan	Ham	423	77
	Spam	77	423	500

Berdasarkan hasil pada Tabel 7. diketahui bahwa komentar ham yang salah diklasifikasikan sebagai komentar spam ada sebanyak 77 dari total 500 komentar ham. Sedangkan untuk komentar spam yang salah diklasifikasikan sebagai ham ada sebanyak 77 dari total 500 komentar spam. Dari hasil pembahasan pada paragraf sebelumnya dapat diketahui bahwa metode SVM dengan kernel linier memberikan hasil yang terbaik dengan error terkecil yaitu 2.8% dan ketepatan klasifikasi sebesar 97.33%.

Nilai pada hasil tersebut merupakan nilai yang didapat dari rata-rata keseluruhan semua fold tetapi seharusnya untuk menghitung *hyperplane* baru diperlukan model dari salah satu fold yang memberikan hasil paling optimum sehingga pada Tabel 7 akan ditunjukkan hasil perhitungan klasifikasi tiap fold-nya. Sehingga dapat diketahui fold yang memberikan hasil paling optimum dan model terbaik untuk menghitung *hyperplane* baru. Perlu diketahui bahwa setiap fold memiliki jumlah

data sebanyak 14 komentar dengan rincian 2 komentar spam dan 12 komentar ham.

Dari 70 data komentar pada metode SVM linier terdapat 5 komentar yang mengalami misklasifikasi dengan rincian dari 21 komentar ham ada 7 komentar ham yang salah diklasifikasikan sebagai komentar spam dan dari 32 komentar spam ada 4 komentar spam yang salah diklasifikasikan sebagai komentar ham.

#### 4. KESIMPULAN

Berdasarkan hasil diketahui bahwa komentar ham yang salah diklasifikasikan sebagai komentar spam ada sebanyak 77 dari total 500 komentar ham. Sedangkan untuk komentar spam yang salah diklasifikasikan sebagai ham ada sebanyak 77 dari total 500 komentar spam.

Nilai pada hasil tersebut merupakan nilai yang didapat dari rata-rata keseluruhan semua fold tetapi seharusnya untuk menghitung *hyperplane* baru diperlukan model dari salah satu fold yang memberikan hasil paling optimum sehingga dapat diketahui fold yang memberikan hasil paling optimum dan model terbaik untuk menghitung *hyperplane* baru. Perlu diketahui bahwa setiap fold memiliki jumlah data sebanyak 14 komentar dengan rincian 2 komentar spam dan 12 komentar ham.

Dari 70 data komentar pada metode SVM linier terdapat 5 komentar yang mengalami misklasifikasi dengan rincian dari 21 komentar ham ada 7 komentar ham yang salah diklasifikasikan sebagai komentar spam dan dari 32 komentar spam ada 4 komentar spam yang salah diklasifikasikan sebagai komentar ham.

#### 5. SARAN

Dari penelitian yang telah dilakukan masih terdapat kekurangan yang dapat diperbaiki oleh pengembang berikutnya. Pengembang berikutnya dapat menggunakan metode preprocessing dan ekstraksi ciri yang berbeda untuk mendapatkan nilai masukan yang lebih baik untuk digunakan pada proses klasifikasi *Support Vector Machine* (SVM).

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang tidak terhingga kepada para teman-teman mahasiswa di lingkungan Universitas ARS, juga kepada dosen dan pengurus baik di tingkat Rektorat, Fakultas, Jurusan maupun Lembaga yang telah banyak membantu pelaksanaan penelitian ini.

#### DAFTAR PUSTAKA

- [1] Difika, Fifit, Dakwah melalui Instagram (studi analisis materi dakwah dalam Instagram Yusuf Mansur, Felix Siau, Aa Gym, Arifin Ilham). Undergraduate (S1) thesis, UIN Walisongo, 2016.
- [2] YR Sipayung, R Sulistyowati - Multimatrix, 2020.
- [3] Nur Azizul Haqimi, Dr. Nur Rokhman, S.Si., M.Kom., Dr. Sigit Priyanta, S.Si., M.Kom. DETEKSI Komentar Spam Pada Instagram Menggunakan Complementary Naive Bayes. 2019.
- [4] Lidya, S. K., Sitompul, O. S., & Efendi, S., Sentiment Analysis Pada Teks Bahasa Indonesia Menggunakan Support Vector Machine (SVM). Seminar Nasional Teknologi Dan Komunikasi, 2015.
- [5] Pratama, A., Wihandika, R. C., & Ratnawati, D. E. Implementasi Algoritme Support Vector Machine (SVM) untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer E-ISSN, 2548, 964X, 2(4), 1704–1708, 2018.