



## Identifikasi Komentar Spam Pada Sosial Media

**Zena Lusi**

Institut Teknologi dan Bisnis Nasional

Email: [zena\\_lusi@itbn.ac.id](mailto:zena_lusi@itbn.ac.id)

**Ayu Eka Saputri**

Universitas Bina Darma

Email: [aesptr22@gmail.com](mailto:aesptr22@gmail.com)

**Tri Basuki Kurniawan**

Universitas Bina Darma

Email: [tribasukikurniawan@binadarma.ac.id](mailto:tribasukikurniawan@binadarma.ac.id)

**Abstract:** *The use of social media is already powerful and difficult to avoid. Social media users are not only limited to the general public, but also public figures and even economic actors who use social media as a means of marketing. In every post from the account owner, there will always be followers who can give likes and comments. Unfortunately, not all comments are related to the uploaded post. One of the most annoying comments is spam comments. Spam comments are comments that are not clear and contain about business (promos / selling), links or various other things that are promoting something. Using the Naive Bayes algorithm, this study wants to identify spam comments, especially on Instagram social media. Where the data is retrieved using the tools provided by Google. Which is then processed with the Rapidminer application to get the Naive Bayes calculation results.*

**Keywords:** *Data Mining, Naive Bayes, Rapidminer, Instagram.*

**Abstrak:** *Penggunaan media sosial sudah sangat berpengaruh dan sulit untuk dihindari. Pengguna sosial media tidak hanya terbatas sebagai masyarakat umum, tetapi juga public figure dan bahkan para pelaku ekonomi yang memanfaatkan media sosial sebagai sarana pemasaran. Dalam setiap postingan dari para pemilik akun, akan selalu ada para follower (pengikut) yang dapat memberikan like dan juga komentar. Sayangnya, tidak semua komentar berkaitan tentang postingan yang diunggah tersebut. Salah satu komentar yang sangat mengganggu adalah komentar spam. Komentar spam adalah komentar yang tidak jelas dan berisikan tentang bisnis (promo/berjualan), link atau berbagai hal lainnya yang bersifat mempromosikan sesuatu. Menggunakan algoritma Naive Bayes, penelitian ini ingin mengidentifikasi komentar spam khususnya pada media sosial instagram. Dimana data di ambil menggunakan tools yang telah disediakan oleh google. Yang kemudian di olah dengan aplikasi rapidminer untuk mendapatkan hasil perhitungan Naive Bayes.*

**Kata kunci:** *Data Mining, Naive Bayes, Rapidminer, Instagram*

### PENDAHULUAN

Dikutip dari halaman website Sekretariat Kabinet Republik Indonesia tahun 2016 ([setkab.go.id](http://setkab.go.id)), Presiden Joko Widodo mengungkapkan bahwa ada sekitar 132 juta pengguna internet yang aktif atau sekitar 52% dari jumlah penduduk yang ada di Indonesia. Sekitar 129 juta memiliki akun media sosial dan rata-rata menghabiskan waktu 3,5 jam per hari hanya untuk konsumsi internet melalui sebuah ponsel. Artinya, penggunaan media sosial sudah sangat berpengaruh dan sulit untuk dihindari.

Dengan adanya sosial media, banyak orang selalu tidak tahan untuk memposting tentang kehidupan sehari-hari mereka. Baik dengan tujuan untuk pamer, atau hanya sekedar ingin berbagi. Karena itulah, pengguna sosial media tidak hanya terbatas sebagai masyarakat umum, tetapi juga *public figure* dan bahkan para pelaku ekonomi yang memanfaatkan media sosial sebagai sarana pemasaran. Dalam setiap posting dari para pemilik akun, akan selalu ada para *follower* (pengikut) yang dapat memberikan *like* dan juga komentar. Pada setiap *like* dan komentar itu, akan menambahkan popularitas sebuah postingan sehingga pemilik postingan akan merasakan sensasi tersendiri karena postingannya mendapatkan banyak tanggapan.

Namun sayangnya, tidak semua komentar berkaitan tentang postingan yang diunggah tersebut. Salah satu komentar yang sangat mengganggu adalah komentar spam. Komentar spam adalah komentar yang tidak jelas dan berisikan tentang bisnis (promo/berjualan), link atau berbagai hal lainnya yang bersifat mempromosikan sesuatu. Komentar spam adalah komentar yang mengganggu dan berulang-ulang.

Dalam penelitian ini, data Instagram yang akan digunakan dalam penelitian ini adalah data artis yang memiliki jumlah pengikut lebih dari 10 juta. Pada akun tersebut akan diambil 30 status terbaru dan 50 komentar terbaru. *Dataset* yang terbentuk akan dilakukan pelabelan secara manual menggunakan tenaga ahli untuk dapat digunakan sebagai data latih sistem *supervised learning* deteksi komentar spam menggunakan Naive Bayes.

## METODOLOGI PENELITIAN

Penelitian ini menggunakan Metode Penelitian Deskriptif. Dimana metode ini memiliki tujuan untuk menjelaskan suatu keadaan atau fenomena yang sedang terjadi saat ini menggunakan prosedur ilmiah dalam menjawab masalah secara aktual, (Sugiyono, 2011). Penelitian ini akan memberitahukan fenomena yang terjadi di masyarakat melalui media sosial Instagram. Menggunakan beberapa *tools* yang telah disediakan oleh *google*, penelitian ini akan menjelaskan bagaimana data akan diambil, disimpan, dan diolah hingga menghasilkan sebuah informasi.

### Metode Pengumpulan Data

Pengumpulan data dilakukan untuk mengumpulkan data yang berhubungan dengan penelitian yang akan dilakukan. Tahap ini, Pengambilan data diambil menggunakan beberapa *tools*, yakni *Spreadsheet*, *Exportcomment* dan *Spatula*. Selain itu metode



dalam universe saling bebas dan tidak berhubungan satu sama lain. Dalam kaitannya dengan data latih, setiap data latih memiliki atribut dan satu untuk label kelas, maka kemungkinan suatu data baru masuk ke dalam suatu kelas dapat didefinisikan dengan persamaan berikut :

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

**Gambar 2.1** Rumus Algoritma *Naive Bayes*

Keterangan :

- X = Data *class* yang belum diketahui.  
 H = Hipotesis data X yang merupakan suatu *class* spesifik.  
 P (H|X) = Probabilitas hipotesis H berdasarkan kondisi X (*posteriori probability*).  
 P (H) = Probabilitas hipotesis H (*prior probability*).  
 P (X|H) = Probabilitas X berdasarkan kondisi pada hipotesis H.  
 P(X) = Probabilitas dari X.

Dimana X adalah bukti, H adalah hipotesis, P(H|X) adalah probabilitas bahwa hipotesis H benar untuk bukti X atau dengankata lain P(H|X) merupakan probabilitas posterior H dengan syarat X, P(X|H) adalah probabilitas bahwa bukti X benar untuk hipotesis H atau probabilitas posterior X dengan syarat H, P(H) adalah probabilitas *prior* hipotesis H, dan P(X) adalah probabilitas *prior* bukti X.

Algoritma *Naive Bayes* memiliki beberapa tahapan untuk mendapatkan output, yaitu :

1. Menentukan pengelompokkan untuk kriteria masing-masing.
2. Menentukan matriks data untuk data yang sudah mengalami pengelompokkan data sesuai kriteria. (Suyanto, 2017:126-127).

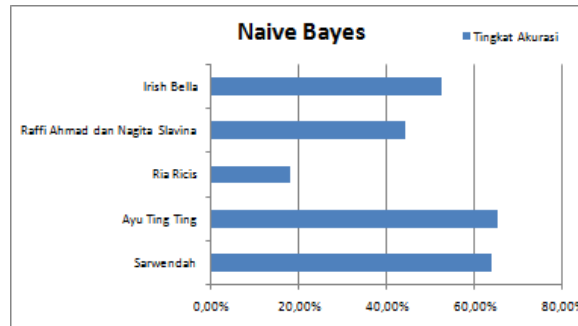
## HASIL

Seperti yang telah di jelaskan sebelumnya, maka hasil dari penelitian ini adalah untuk mengklasifikasikan data komentar spam pada Instagram. Data tersebut akan dianalisis dan di visualisasi menjadi suatu informasi.

Berikut tampilan hasil keseluruhan dari masing-masing akun instagram yang telah di proses melalui *Naive Bayes* tersebut:

Akun	Tingkat Akurasi	Class Precision		Class Recall	
		Spam	Bukan Spam	True Spam	True Bukan Spam
Sarwendah	63.81%	33.04%	97.76%	94.18%	57.04%
Ayu Ting Ting	65.04%	55.12%	94.68%	96.87%	41.40%
Ria Ricis	18.01%	3.61%	99.56%	97.87%	15.43%
Raffi Ahmad dan Nagita Slavina	44.22%	27.01%	97.80%	97.45%	30.09%
Irish Bella	52.43%	20.74%	99.67%	98.94%	45.77%

**Tabel 1** Tabel Hasil Keseluruhan.



**Gambar 2** Grafik Naive Bayes.

## SIMPULAN

Dari penelitian yang telah dilakukan terhadap spam komentar pada instagram, dapat disimpulkan bahwa:

1. Penarikan data pada media sosial instagram telah dilakukan menggunakan *tools Spreadsheet, Spatula* dan *Exportcomment*. Pada proses ini tidak diperlukan *Authentication* terlebih dahulu untuk bisa mengakses instagram API, karena pada tahap ini penarikan data langsung menggunakan *spreadsheet, spatula* dan *exportcomment*.
2. Dari hasil yang telah didapatkan menggunakan Rapidminer dapat disimpulkan bahwa akurasi yang terbesar dari kelima akun instagram tersebut adalah 65.04% spam dari akun milik Ayu Ting Ting.

## DAFTAR RUJUKAN

- Aditiya, S. 2015. InstaInsta Instagram. <http://sehabiskelas.blogspot.com/2015/01/insta-insta-instagram.html>. Jum'at, 16 Januari 2019.
- Chrismanto, A. R. and Lukito, Y. 2017. *Deteksi Komentar Spam Bahasa Indonesia pada Instagram menggunakan Naive Bayes*, pp.50-58.
- Chrismanto, A. R. and Lukito, Y. 2017. *Identifikasi Komentar Spam pada Instagram*.
- Destuardi, I. and Sumpeno, S., 2009. Klasifikasi Emosi Untuk Teks Bahasa Indonesia Menggunakan Metode Naive Bayes. *Seminar Nasional Pascasarjana Institut*

*Teknologi Sepuluh Nopember,*

- Jaelani, I. Data Mining. <https://ocw.upj.ac.id/files/Handout-TIF311-DM-1.pdf>. 12 November 2019.
- Kominfo.go.id.2017. Pemerintah Ingin Media Sosial di manfaatkan untuk hal yang Produktif. [https://www.kominfo.go.id/content/detail/8637/pemerintah-ingin-media-sosial-dimanfaatkan-untuk-hal-produktif/0/sorotan\\_media](https://www.kominfo.go.id/content/detail/8637/pemerintah-ingin-media-sosial-dimanfaatkan-untuk-hal-produktif/0/sorotan_media). 12 November 2019.
- Mkhuda.2014. Mengintip Dapur Pacu Instagram. <https://mkhuda.com/teknologi/mengintip-dapur-pacu-instagram/>. 12 November 2019.
- Pradhika Yuda. Spamming. <https://www.academia.edu/19628032/spamming>.
- Samodra, J., Sumpeno, Surya., Hariadi, M. 2009. *Klasifikasi Dokumen Teks Berbahasa Indonesia dengan menggunakan Naive Bayes*.
- Suyanto, 2017. *Data Mining untuk Klasifikasi dan Klusterisasi data*. Bandung : Informatika Bandung.
- Widodo, A. W. and Mahmudy, W. F. 2014. *Klasifikasi Artikel Berita Secara Otomatis menggunakan Metode Naive Bayes Classifier yang di modifikasi*.