

# ANALISIS SENTIMEN MEDIA SOSIAL UNIVERSITAS AMIKOM YOGYAKARTA SEBAGAI SARANA PENYEBARAN INFORMASI MENGGUNAKAN ALGORITMA KLASIFIKASI SVM

Muhammad Akbar Maulana<sup>1)</sup>, Arief Setyanto<sup>2)</sup>, Mei P. Kurniawan<sup>3)</sup>

<sup>1)</sup> Magister Teknik Informatika, Universitas AMIKOM Yogyakarta

<sup>2)</sup> Fakultas Sains dan Teknologi, Universitas AMIKOM Yogyakarta

<sup>3)</sup> Fakultas Ekonomi dan Sosial, Universitas AMIKOM Yogyakarta

Jl. Ring Road Utar, Condong Catur, Depok, Sleman, Yogyakarta 55281

Email : muhammad.0874@students.amikom.ac.id<sup>1)</sup>, arief\_s@amikom.ac.id<sup>2)</sup>, emailketiga@gmail.com<sup>3)</sup>

## Abstrak

Sosial media adalah media hiburan yang sangat menyita waktu dan perhatian para penggunanya. Tidak hanya hiburan semata yang disajikan oleh sosial media tetapi juga berbagai informasi penting seputar politik, agama maupun penjualan produk. Indonesia mulai memasuki masa dimana e-commerce mulai banyak digunakan.

Peran pemasaran sebuah produk atau menyebarkan informasi sangatlah penting untuk meningkatkan minat konsumen dan pengetahuan akan suatu hal. Banyak cara bisa dilakukan termasuk dengan media dunia maya menggunakan facebook atau twitter. Universitas Amikom Yogyakarta adalah salah satu perguruan tinggi yang mulai merubah statusnya dari STMIK menuju Universitas.

Dengan menerapkan analisis sentimen dan menggunakan algoritma klasifikasi SVM dapat diketahui bagaimana hasil feedback dari pemasaran atau penyebaran informasi yang didapatkan dari konsumen, masyarakat bahkan instansi tertentu.

**Kata kunci:** Sosial Media, informasi, sentimen analisis, SVM.

## 1. Pendahuluan

Era media digital mulai masuk dengan pesat ke dalam pemasaran di Indonesia, tidak hanya dari media offline saja tetapi media online turut memberikan kontribusinya dalam menyampaikan informasi kepada para pencari informasi [1]. Sosial media merupakan salah satu media yang mempermudah komunikasi interaktif antara pengusaha dengan siapapun, termasuk konsumen, penyalur, pemasok dan berbagai pihak yang berkepentingan; kapanpun dan dimanapun berada [2]. Sosial network menduduki peringkat pertama dalam penyebaran konten, dan 85.4% penduduk Indonesia selalu mengurus profil sosial media mereka [3].

Banyak peluang yang dapat dimanfaatkan dari hal tersebut salah satunya yaitu menyebarkan informasi perubahan status instansi Amikom Yogyakarta. *Feedback* dari masyarakat sangatlah penting agar mengerti bagaimana perubahan ini di mata masyarakat. Untuk mendapatkan hasil akurat dibutuhkan metode dan algoritma yang tepat, agar tidak membuang waktu pada saat mengumpulkan dan menganalisis data tersebut. Maka dari itu dibutuhkan API dari sebuah platform sosial media untuk mengambil data, setelah mendapatkan data maka dilanjutkan dengan analisis sentiment yang berguna untuk menganalisis komentar-komentar di twitter tersebut untuk kemudian diterjemahkan menjadi suatu yang lebih bermakna [4]. Analisis ini juga berguna untuk berbagai masalah spesifik bagi para praktisi dan peneliti dibidang *human-computer interaction*, serta mereka yang berasal dari bidang sosiologi, pemasaran dan periklanan, psikologi, ekonomi dan politik [5].

Tujuan penelitian ini dapat mengidentifikasi sosial media mana yang tepat untuk melakukan penyebaran informasi agar pihak marketing akan lebih fokus untuk menekan social tersebut. Penelitian menerapkan pengambilan data yang besar atau bisa disebut dengan big data dan mengolah data tersebut menggunakan analisis sentiment.

## Sentimen Analisis

Analisis sentiment disebut juga *opinion mining*, adalah bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi orang-orang terhadap entitas seperti produk, layanan, organisasi, individu, isu, peristiwa, topik, dan atribut mereka. Ini merupakan ruang masalah besar [6]. Hasil akhir yang akan dicapai pada analisis ini menjadi suatu alat otomatis untuk memberikan informasi berupa sekumpulan teks, untuk digunakan pada orang yang mengambil keputusan atau yang memberi keputusan.

Saat ini, analisis sentimen telah mendapatkan nilai lebih dengan munculnya jejaring sosial. Difusi besar dan peran mereka dalam masyarakat modern merupakan salah satu hal baru yang paling menarik dalam beberapa tahun

terakhir, menarik minat para peneliti, jurnalis, perusahaan, dan pemerintah (Pozzi, Fersini, Messina, & [7].

### Text Mining

Text mining merupakan metode klasifikasi yang merupakan variasi dari data mining berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar (Kurniawan, Effendi, & Sitompul, 2012)[8]. Text mining adalah salah satu teknik yang dapat digunakan untuk melakukan klasifikasi dimana, text mining merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar [9]. Selain klasifikasi, text mining juga digunakan untuk menangani masalah clustering, information extraction, dan information retrieval [10].

Pada Bahasa R untuk memining tulisan yang telah kita kumpulkan, dengan cara menginstall tm package dan menanggil library tm yang sebelumnya di install pada Rstudio.

```
# text mining
library("tm")
library("Rcurl")

docs <- Corpus(VectorSource(message))

toSpace <- content_transformer(function(x, pattern) gsub(pattern, " ", x))
docs <- tm_map(docs, toSpace, "/")
docs <- tm_map(docs, toSpace, "@")
docs <- tm_map(docs, toSpace, "\\n")

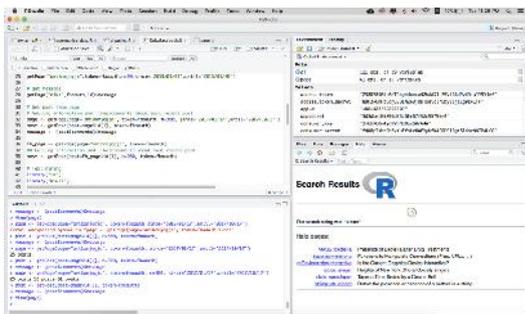
# Convert the text to lower case
docs <- tm_map(docs, content_transformer(tolower))

# Remove numbers
docs <- tm_map(docs, removeNumbers)
```

Gambar 1. Syntax dari Rstudio

### Bahasa R

R bukan saja bahasa tetapi juga lingkungan/environment untuk komputasi statistik dan grafik. R merupakan project GNU yang dikembangkan oleh Bell Laboratories (sebelumnya AT&T, sekarang Lucent Technologies). Team pengembang R adalah John Chamber dan teman-temannya [11]. Sebagian besar sistem R ditulis dalam R, Untuk tugas komputasi intensif, C, C ++, dan kode FORTRAN dapat dihubungkan dan dipanggil pada waktu berjalan [12]. Salah satu aplikasi untuk menjalankan Bahasa R adalah Rstudio. Tool ini memiliki interface yang mudah dimengerti.



Gambar 2. GUI Rstudio

### RFacebook

Facebook adalah situs web tempat Anda terhubung dan berbagi dengan teman. Dan sama seperti ada banyak cara berbeda untuk berinteraksi dengan teman di dunia nyata, ada banyak cara untuk melakukannya di Facebook [13].

Rfacebook adalah sebuah paket pada bahasa R yang menyediakan fungsi untuk pengguna dapat mengakses API facebook untuk mendapatkan informasi tentang semua yang ada di facebook seperti: posts, public, pages, groups, bahkan data pribadi dari seseorang yang dituju. Untuk saat ini CRAN yang dipakai menggunakan versi 0.6.15.

Untuk mengakses API facebook ada beberapa ketentuan khusus, yaitu membuat permintaan untuk mengakses data tersebut, salah satu caranya dengan menggunakan token untuk mengisi appid dan appsecret dengan cara mengakses <https://developers.facebook.com/apps>



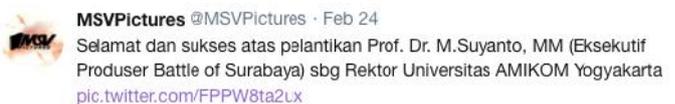
Gambar 3. Data Mentah Facebook

### RTwitter

Twitter adalah salah satu situs jejaring sosial sama halnya dengan facebook. Twitter dibuat oleh Jack Dorsey di tahun 2006 dan di tahun itu juga muncul di web dan diberi alamat <http://www.twitter.com> hingga saat ini.

Rtwitter adalah sebuah paket sama seperti dengan Rfacebook yang berfungsi sebagai perantara untuk pengguna dapat mengakses API twitter dengan mudah untuk mendapatkan informasi berupa tweet, retweet, dll. CRAN twitter lebih unggul dibanding facebook dengan lebih banyak syntax didalamnya.

Berbeda dengan facebook untuk mendapatkan akses, pengguna menginputkan 4 token untuk mendapatkan akses informasi yang diinginkan, yaitu : consumer key, consumer secret, access token, access token secret.



Gambar 4. Data Mentah Twitter

### SVM

Support Vector Machines (SVM) adalah seperangkat metode pembelajaran yang menganalisis data dan mengenali pola, digunakan untuk klasifikasi dan analisis regresi, tidak hanya itu teknik ini dapat melakukan prediksi dan penilaian tentang sebuah system.

SVM sangat terkenal dengan komunitas klasifikasi pola yang secara default, para periset di bidang ini menggunakannya sebagai penggolongan dasar untuk menetapkan superioritas classifier yang diusulkan oleh mereka [14]. Tujuan dari SVM memberikan nilai dari banyaknya kemunculan sebuah kata atau dapat mengklasifikasi komen positive dan negative.

Pada teknik Support Vector Machine (SVM) akan dikelompokkan 2 data atau lebih dan akan dibuat garis untuk memisahkan kelompok data tersebut. Garis ini dikenal dengan hyperplane. Teknik SVM bertujuan untuk mencari hyperplane yang optimal.

Diketahui training vector  $x_i \in \mathbb{R}^p, i = 1, \dots, n$  dalam dua kelas, dan vector  $y \in \{1, -1\}^n$ , svm menyelesaikan permasalahan berikut

$$\min_{w,b,c} \frac{1}{2} w^t w + c \sum_{i=1}^n c_1 \quad (1)$$

Subject to  $y_i(w^T \phi(x_i) + b) \geq 1 - \mathbb{C}_i$ ,

$\mathbb{C}_i \geq 0, i = 1, \dots, n$

Dengan elemen keduanya adalah

$$\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - e^T \quad (2)$$

Subject to  $y^T \alpha = 0$

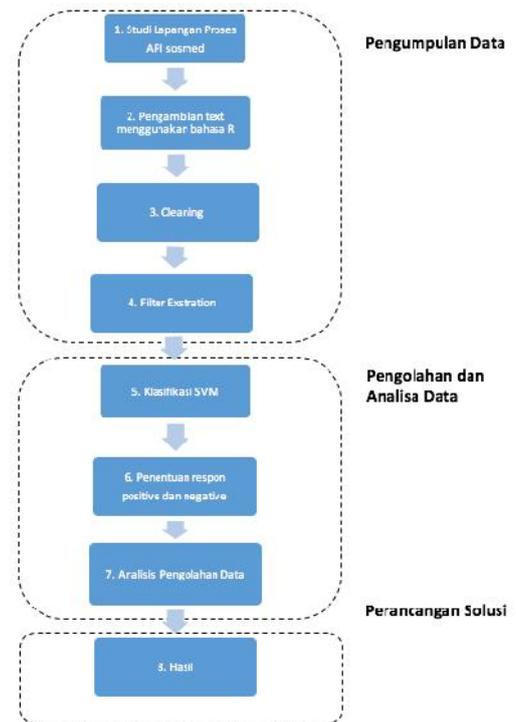
$0 \leq \alpha_i \leq C, i = 1, \dots, n$

Dimana  $e$  adalah vector dari keseluruhan,  $C > 0$

Adalah batas atas,  $\varphi$  adalah  $n$  dari  $n$  positif semidefinite matrix [15].

## 2. Pembahasan

Dalam penelitian ini terdapat beberapa langkah yang digambarkan dalam diagram pada gambar berikut [16]:



Gambar 5. Alur Kerja

Keterangan dari gambar (diatas) sebagai berikut:

1. Melakukan studi observasi penggunaan sosial media yang sudah berjalan di Universitas AMIKOM Yogyakarta.
2. Pemetaan data mengacu pada sentimen analis dengan menggunakan bahasa R. pada tahapan ini dilakukan pengambilan data dari API twitter dan facebook yang sudah didapat untuk memudahkan dalam pengumpulan data.
3. Data yang telah diambil akan di bersihkan dahulu untuk mengilangkan symbol dan tanda baca.
4. Setelah data bersih maka langkah selanjutnya yaitu memfilter data tersebut untuk melihat seberapa banyak kata yang dipakai, disini juka akan dilakukan tokenizing, stopwords removal dan stemming
5. Pengolahan data menggunakan klasifikasi SVM.
6. Hasil dari SVM akan dikelompokkan berdasarkan respon positive negative dan netral.
7. Analisis dari hasil pengolahan data yang berupa feedback yang telah dikumpulkan
8. Penyusunan rekomendasi penggunaan sosial berdasarkan analisis sentimen dan klasifikasi SVM.

Proses analisis dengan algoritma SVM menggunakan 2 data yang berbeda yaitu twitter dan facebook. Dari sosial media tersebut diambil masing-masing 300 data yang akan dijadikan data uji. Tema data yang diambil yaitu Universitas Amikom Yogyakarta dan penerimaan mahasiswa baru, sehingga keyword yang digunakan yaitu berhubungan dengan kata tersebut. Untuk pengambilan data menggunakan bahasa inggris dan

bahasa indonesia. Setelah data didapat maka memisahkan kata yang berunsur positif, negatif, dan netral. maka didapatkan jumlah data hasil seperti pada tabel dibawah ini

**Tabel 1. Jumlah Hasil Sentimen Analisis**

No	Keyword	Jumlah data keseluruhan	Jumlah data uji positif	Jumlah data uji negatif	Jumlah data uji netral
1	@amikomjogja	300	283	2	15
2	#amikomjogja	300	286	5	9
3	Amikom	300	294	3	3
4	Amikomjogja	300	282	3	15
5	Amikom_penerimaan	300	291	7	2
6	amikom	300	265	5	30

Dari hasil percobaan tabel 1 maka dengan hasil data uji positif yang lebih tinggi maka hasil analisis akan menunjukkan sebagai opini positif. Begitu juga sebaliknya, jika data uji negatif lebih tinggi hasil analisis akan menunjukkan opini negatif.

**Tabel 2. Persentase Analisis Sentimen SVM Facebook**

No	Keyword	Positif (%)	Negatif (%)	Netral (%)
1	amikomjogja	94.34%	0.66%	5%
2	Amikom_penerimaan	95.43%	1.57%	3%
3	Amikom	98.10%	0.9%	1%

**Tabel 3. Persentase Analisis Sentimen SVM Twitter**

No	Keyword	Positif (%)	Negatif (%)	Netral (%)
1	@amikomjogja	94.20%	0.8%	5%
2	#amikomjogja	97%	2.3%	0.7%
3	Amikom	88,4%	1.6%	10%

Setelah melihat tabel pertama maka didapatkan hasil persentase prediksi pada tabel 2 dan tabel 3 dari kedua sosial media tersebut, dapat dilihat dengan mengambil banyak data latih yang sama dengan menggunakan 3 keyword. Pada keyword pertama menggunakan keyword yang sama didapatkan facebook memiliki komentar

positif yang lebih besar daripada twitter. Begitu juga dengan keyword kedua dan ketiga nilai positif lebih banyak di pegang oleh facebook.

### 3. Kesimpulan

Adapun kesimpulan yang telah dicapai dari penelitian ini yaitu:

1. Dengan menerapkan implementasi algoritma klasifikasi SVM pada sosial media yang digunakan oleh Universitas Amikom Yogyakarta dapat memprediksi performa dari kedua sosial media yang dimiliki universitas tersebut yaitu facebook dan twitter. Dari hasil ini didapat bahwa facebook memiliki presentase lebih tinggi dari twitter dengan melihat banyaknya feedback dan komentar positif yang didapatkan dari user.
2. Dengan melihat alur yang telah dilakukan ada kemungkinan bahwa metode ini dapat diterapkan untuk universitas atau departemen lain.
3. Melihat dari perbedaan data yang didapat dari facebook dan twitter tidak mempengaruhi kecepatan hasil akhir dari pengelompokan nilai positif, negatif dan netral.

### 4. Saran

Adapun saran untuk penelitian setelah ini yaitu:

1. Untuk mendapatkan hasil yang lebih tepat dapat dilakukan survey langsung untuk kecocokan antara metode yang dipakai dengan dunia nyata.
2. Karena telah mendapatkan data yang cukup kedepannya akan dicoba untuk menggunakan algoritma lain untuk melihat algoritma mana yang lebih tepat dan cepat untuk melakukan prediksi pada analisis sentiment.

### Daftar Pustaka

- [1] Azkia, M. R., Candra, B. P., Sadida, R., Rezeki, N., & Rendy, M. C. (2017, Februari 4). Perancangan Sistem Analisis Sentimen Masyarakat Pada Sosial Media Dan Portal Berita. *Seminar Nasional Teknologi Informasi dan Multimedia 2017*, 199-204.
- [2] Priambada, S. (2015, 11 2-3). Manfaat Penggunaan Media Sosial Pada Usaha Kecil Menengah (Ukm). 41-46.
- [3] Louie, J. (2014). *wave 8 : the language of content*. Retrieved from wave.umww.com: <http://wave.umww.com/assets/pdf/wave%5Ctextunderscore%5Cn8-the-language-of-content.pdf>
- [4] Monarizqa, N., Nugroho, L. E., & Hantono, B. S. (2014). Penerapan Analisis Sentimen Pada Twitter Berbahasa Indonesia Sebagai Pemberi Rating. *Jurnal Penelitian Teknik Elektro dan Teknologi Informasi*, 151-155.
- [5] Hutto, C. J. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Eighth International AAAI Conference on Weblogs and ...*, 216-225.
- [6] Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. (U. o. Graeme Hirst, Ed.) Morgan & Claypool Publisher.

- [7] Pozzi, F. A., Fersini, E., Messina, E., & Liu, B. (2016). *Sentiment Analysis in Social Networks*. Morgan Kaufmann.
- [8] Kurniawan, B., Effendi, S., & Sitompul, O. S. (2012). Klasifikasi Konten Berita Dengan Metode Text Mining. *I* (1), 14-19.
- [9] Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook : Advanced Approaches in Analyzing Unstructured Data*. New York: Cambridge University Press.
- [10] Berry, M., & Kogan, J. (2010). *Text Mining Application and theory*. United Kingdom: WILEY.
- [11] Faisal, M. R. (2017). *Seri Belajar Data Science: Klasifikasi dengan Bahasa Pemrograman R*. Indonesia .Net Developer Community.
- [12] Janoušek, V., Moyen, J.-F., Martin, H., Erban, V., & Farrow, C. (2015). *Geochemical Modelling of Igneous Processes – Principles And Recipes in R Language: Bringing the Power of R to a Geochemical Community*. Springer Geochemistry.
- [13] Abram, C. (2016). *Facebook For Dummies: Edition 6* (Vol. 6). John Wiley & Sons.
- [14] Murty, M., & Raghava, R. (2016). *Support Vector Machines and Perceptrons: Learning, Optimization, Classification, and Application to Social Networks*. Springer.
- [15] Laksana, E. A., & Sulianta, F. (2017, februari 4). Analisis Dan Studi Komparatif Algoritma Klasifikasi Genre Musik. 67-72.
- [16] Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04*, 168-177.

### **Biodata Penulis**

**Muhammad Akbar Maulana**, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Teknik Informatika STMIK AMIKOM Yogyakarta, lulus tahun 2016. Saat ini sedang melanjutkan jenjang Magister Teknik Informatika dan menjadi Staff operasional di Inkubator Amikom Business Park.

**Arief Setyanto**, memperoleh gelar Sarjana Sistem Informasi (S.Si), Jurusan Elektronika dan instrumentasi Universitas Gajah Mada Yogyakarta, lulus tahun 1998. Memperoleh gelar Magister Teknik (MT) Program Pasca Sarjana Magister Teknik Elektro Universitas Gajah Mada Yogyakarta, lulus tahun 2003. Memperoleh gelar Doctor of Philosophy (Ph.D) Program doktoral Scholl of Computer Science and Electronics Engineering, The University of Essex, lulus tahun 2016. Saat ini menjadi Dosen di Universitas AMIKOM Yogyakarta.

**Mei P Kurniawan**, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Sistem Informasi STMIK AMIKOM Yogyakarta, lulus tahun 2008. Memperoleh gelar Magister Teknik Informatika STMIK Amikom Yogyakarta, lulus tahun 2011. Saat ini menjadi sekretaris prodi S1SI, Dosen dan ketua KEJALA (Keluarga Jaringan Alumni AMIKOM) di Universitas Amikom Yogyakarta.

