

# Opinion Analytics dengan Algoritma Rocchio Terhadap Evaluasi *Project-Based Learning* pada Mata Kuliah Teknik Multimedia

Cian Ramadhona Hassolthine<sup>#1</sup>, Riad Sahara<sup>#2</sup>, Ajang Rahmat<sup>#3</sup>

<sup>#123</sup>Program Studi Informatika PJJ, Universitas Siber Asia  
Jl. Harsono RM Ragunan, Pasar Minggu, South Jakarta City

<sup>1</sup>cianhassolthine@lecturer.unsia.ac.id

<sup>2</sup>riadsahara@lecturer.unsia.ac.id

<sup>3</sup>ajang.rahmat0010@student.unsia.ac.id

**Abstrak** — *Opinion analytics* merupakan merupakan salah satu bidang dari *Natural Language Processing* (NLP) mengekstraksi opini dalam bentuk teks. Informasi sentimen dalam suatu kalimat opini dapat dianalisis untuk melihat kecenderungan opini bersifat negatif atau positif. Pada penelitian ini, dataset diambil dari opini mahasiswa melalui sosial media twitter. Total dataset yang terkumpul sebanyak 769 opini. Algoritme yang digunakan adalah Algoritme Hasil dari klasifikasi terdapat tiga kelas opini, yaitu kelas positif Rocchio sebagai klasifikasi (kelas positif = 295, kelas negatif = 298 dan kelas netral = 176). *Output* pada penelitian ini ditentukan dari rubrik penilaian untuk setiap kata yang telah diklasifikasikan. Hasil penelitian ini menunjukkan bahwa model mampu memberikan rekomendasi untuk meningkatkan pelaksanaan penilaian tugas proyek mata kuliah teknik multimedia.

**Kata kunci** — Algoritme Rocchio, Klasifikasi, *Opinion analytics*, *Text mining*, Twitter

## A. Pendahuluan

*Opinion analytics* merupakan merupakan salah satu bidang dari *Natural Language Processing* (NLP) mengekstraksi opini dalam bentuk teks dan mendapatkan informasi sentimen dalam suatu kalimat opini yang dianalisis untuk melihat kecenderungan opini bersifat negatif atau positif [1]. Opini pada kasus-kasus tertentu sering digunakan sebagai data untuk penentuan keputusan terhadap perkembangan sebuah isu yang berkembang. Maraknya penggunaan media sosial di era digital ini, telah ikut mempengaruhi setiap perkembangan isu yang beredar ditengah kehidupan sosial masyarakat. Hal ini dikarenakan media sosial merupakan tempat yang memungkinkan bagi semua orang untuk mengekspresikan opini mereka ke ruang publik [2]. Berdasarkan data Kementerian Komunikasi dan Informatika Republik Indonesia (Kemenkominfo RI) mengungkapkan bahwa pengguna internet dewasa ini sebanyak 63 juta orang, dengan 95 persennya merupakan pengguna jejaring media sosial dan 95 persennya pengguna media twitter. Menurut websindo bahwa media sosial twitter merupakan media sosial dengan pengguna

terbesar keenam setelah youtube, whatsapp, instagram dan line. Banyaknya pengguna twitter dikarenakan selain mudah melakukan penyampaian opini terhadap setiap masalah, juga pada twitter setiap user dapat berupa dapat memanfaatkan data opini sebagai bahan penilaian, tingkat kepuasan dan evaluasi pada kasus yang ingin diselesaikan [3]. Berdasarkan hal tersebut sehingga mendorong beberapa instansi, lembaga maupun beberapa kelompok organisasi tertentu untuk mendapatkan suatu informasi terkait isu yang akan dianalisis.

Di Perguruan Tinggi perlu dilakukan peningkatan dengan proses evaluasi. Dengan evaluasi terdapat data yang akurat dan objektif tentang pelaksanaan suatu kegiatan. Pada penilaian Tugas proyek mata kuliah Teknik Multimedia menggunakan bobot penilaian yang besar, hal ini menimbulkan kendala pada mahasiswa. Materi yang terkait dengan Tugas Proyek ini yaitu dari Materi pertama sampai materi 12. Adapun beberapa kendala dalam proses pembelajaran Teknik Multimedia yaitu, kapasitas kelas mahasiswa yang besar, penyampaian materi dari Asisten/Dosen kurang jelas, beban capaian pertemuan dengan materi yang banyak disampaikan sehingga mahasiswa kurang fokus dalam mengerjakan tugas proyek akhir. Penelitian ini diambil dari data teks melalui twitter, karena mahasiswa dapat dengan bebas mengirimkan opini di luar jam praktikum.

Metode penelitian dilakukan menggunakan teknik *text mining*, kemudian dihitung pembobotan di setiap kata dengan algoritme *term frequency-invers document frequency* (TF-IDF) dan dilakukan klasifikasi dengan Algoritme Rocchio untuk penentuan kelas opini positif, negatif dan netral. *Output* yang ditampilkan berupa rekomendasi dari rubrik penilaian dengan perbandingan klasifikasi opini twitter dengan nilai proyek.

## B. Dasar Teori Mutakhir

Penelitian [4] melakukan analisis sentimen pada opini mahasiswa menggunakan *Natural Language Processing* (NLP). Penelitian tersebut membahas *opinion mining* dari data mahasiswa Ilmu Komputer sebanyak 150 opini. Algoritme

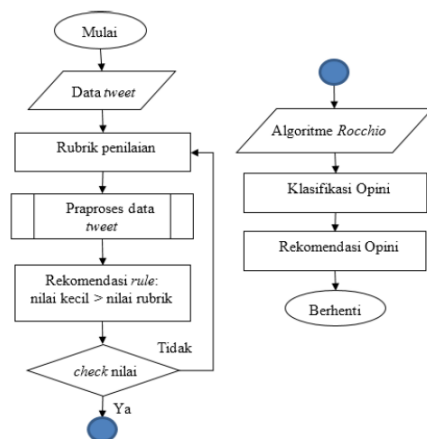
yang digunakan yaitu *Naïve Bayes Classification* untuk proses klasifikasi dan analisis opini mahasiswa.

Penelitian [5] melakukan analisis tentang *project-based learning* dengan judul *examining learning transformation in project-based learning process*, mengadopsi teori perubahan *Lewin (unfreezing, movement, dan refreezing)* sebagai acuan analitik untuk membahas kurikulum pembelajaran 2 berbasis proyek di Universitas UNI-X.

Penelitian [6] melakukan penelitian mengenai analitik data dalam pembelajaran yang bertujuan untuk memperkenalkan pembaca ke *Opinion Mining*, dan menyajikan strategi untuk menghitung polaritas sentimen komentar siswa.

### C. Metode Penelitian

Penelitian ini melalui beberapa tahapan yaitu pengambilan data *tweet*, rubrik penilaian, praproses data *tweet*, rekomendasi *rule* dan klasifikasi opini yang akan menghasilkan rekomendasi opini. Tahapan penelitian dapat dilihat pada Gambar 1.



GAMBAR 1. TAHAPAN PENELITIAN

### D. Hasil dan Analisa

Data yang digunakan pada penelitian ini adalah data teks dari media sosial Twitter. Proses pengambilan data ini dilakukan selama dimulai perkuliahan. Setiap mahasiswa mengirimkan *tweet* terkait opini terhadap materi di setiap pertemuan praktikum dengan kata kunci yang telah ditentukan dengan hastag #TEKMULTEK yaitu singkatan dari Teknik Multimedia Tekom, #SVIPB yaitu singkatan dari Sekolah Vokasi IPB dan #TEK2B yaitu singkatan dari Tekom kelas 2B. Penyimpanan data menggunakan program pada bahasa pemrograman R dengan format file .csv (*comma-separated values*) untuk dilakukan proses selanjutnya. Tabel 1 menunjukkan beberapa contoh data *tweet* tentang proyek Teknik Multimedia.

TABEL 1  
SAMPLE DATA TWEET OPINI PROYEK

Username	Tweet
@RahayMutiar	Praktikum minggu kemaren, cuma disuruh lanjutin materi mulmed yang diterapin buat proyek kelompok #TEKMULTEK #SVIPB

	#TEK2B
@Deezcsr	Proyek kelompok saya, desain yang dibikin di flashnya tidak sesuai dengan storyboard #TEKMULTEK #SVIPB #TEK2B
@Syfaa17	Ternyata banyak banget kekurangannya, dan banyak banget yang harus diperbaiki #TEKMULTEK #SVIPB #TEK2B

Hasil dari Tabel 1 menunjukkan contoh data *tweet* opini proyek. Opini dari masing-masing mahasiswa masih menggunakan kata yang tidak baku. Penggunaan *hastag* dan simbol-simbol lainnya. Dalam hal ini perlu dilakukan praproses data untuk memudahkan dalam pengolahan data teks.

### Praproses

Terdapat tiga tahapan praproses dalam penelitian ini, yaitu: tokenisasi, normalisasi, penghapusan *stopwords*. Pada tahap ini mengubah semua huruf kapital pada menjadi huruf kecil. Kemudian menghapus *mention*, *URL*, tanda baca dan angka menggunakan fungsi *gsub*.

### Tokenisasi

Dari 1497 data *tweet* yang diperoleh, dilakukan proses tokenisasi. Pengecekan *tweet* dari karakter pertama sampai karakter terakhir dengan menghapus titik (.), koma (,), spasi atau tanda pemisah lain (- + / & ! ?), maka digabungkan dengan karakter selanjutnya. Salah satu ilustrasi hasil sebelum dan setelah proses tokenisasi dapat dilihat pada Tabel 2.

TABEL 2  
SAMPLE DATA SEBELUM DAN SETELAH PROSES TOKENISASI

Tweet	Sebelum tokenisasi	Setelah tokenisasi
1	@Agustin42515535 Pada saat mengerjakan animasi, dosen yang menerangkan materi kuliah Komputer Multimedia agak terlalu cepat dalam menerangkan, sehingga saya merasa keteteran.#TEKMULTEK #SVIPB#TEK2B #RANKEDTEKMUL	agustin pada saat mengerjakan animasi dosen yang menerangkan materi kuliah komputer multimedia agak terlalu cepat dalam menerangkan sehingga saya merasa keteteran
2	@dewiratnawati08 Membuat animasi ekspresi ketika mengerjakan proyek disitu saya bingung mau membuat apa dan akhirnya saya membuat animasi bahagia, saya merasa kesulitan karena bingung akan gerakan animasinya #TEKMULTEK#SVIPB	membuat animasi ekspresi ketika mengerjakan proyek disitu saya bingung mau membuat apa dan akhirnya saya membuat animasi bahagia, saya merasa kesulitan karena bingung akan gerakan animasinya

	#TEK2B #RANKEDTEKMUL	
--	-------------------------	--

## Normalisasi

Pada tahap ini dilakukan proses penggantian kata yang tidak baku menjadi kata baku. Kata baku akan cenderung lebih kecil ambiguitas dibandingkan dengan kata yang tidak baku [7]. Setelah hasil tokenisasi diperoleh, proses pembakuan kata dilakukan dengan proses pengubahan kata yang tidak baku menjadi kata baku sesuai [8]. Penggunaan kata *sya*, *ku*, dan *gue* akan diubah menjadi kata *saya*, agar kata-kata tersebut memiliki makna yang sama. Proses pengantian kata tidak baku menjadi baku menggunakan daftar kata yang sudah ada. Hal ini dilakukan untuk memudahkan proses pengantian kata

### Penghapusan *stopwords*

Setelah dilakukan proses normalisasi lalu data tersebut digabungkan ke dalam satu dokumen. Kemudian dilakukan proses penghapusan kata *stopwords* dan dilanjutkan penyusunan *term-document matrix* dengan setiap kata yang diasumsikan sebagai dokumen. Berdasarkan *term-document matrix* dari hasil praproses maka didapatkan *term* yang banyak muncul dapat dilihat pada Gambar 2.



GAMBAR 2. *WORDCLOUD TERM* YANG SERING MUNCUL

*Term Frequency-Inverse Document Frequency*  
(TF-IDF)

Pembentukan bobot tiap kata dari data *tweet* yang terkumpul dengan menggunakan teknik TF-IDF. Sebagai ilustrasi, terdapat dua *tweet* yang sudah bersih sebagai berikut :

1. Awal masuk matkul ini sedikit mengantuk karena terbawa oleh suasana yang ada di luar
2. Dan untuk saat ini belum kepikiran untuk project dari multimedia apa sedikit agak tegang karna pelajarannya belum diajari

Dari data *tweet* tersebut kemudian dipilih kata-kata unik dari dua *tweet* tersebut sehingga menjadi matrik yang terlihat pada Tabel 3.

TABEL 3  
Matriks *term data tweet*

<i>term</i>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>.</b>	<b>.</b>	<b>1498</b>
ada	1	0	0	1	0	0	0
diluar	1	0	0	0	0	0	0
masuk	1	0	0	0	0	0	0
matkul	1	0	0	0	0	0	0

mengantuk	1	0	0	0	0	0	0
suasana	1	0	0	0	0	0	0
terbawa	1	0	0	0	0	0	0
diajari	0	1	0	0	0	0	0
kepipikiran	0	1	0	0	0	0	0
multimedia	0	1	1	1	0	0	1
pelajarannya	0	1	0	0	0	0	0
project	0	1	0	0	0	0	0

Dari matrik tersebut kemudian dilakukan proses perhitungan bobot untuk semua kata, hasil matriks TF-IDF dapat dilihat Tabel 4.

TABEL 4  
Matriks TF-IDF Data *TWEET*

<i>term</i>	<b>bobot</b>
multimedia	141.681.555.015.478
pertemuan	141.957.417.259.386
animasi	152.739.320.208.989
praktikum	17.817.892.969.082
komputer	182.642.145.092.897

## Klasifikasi Algoritme Rocchio

Penentuan kelas positif, negatif dan netral yaitu dengan membandingkan nilai *centroid* (tengah) di setiap kata yang paling tinggi. Pada proses sentimen nalisis dilakukan klasifikasi menggunakan algoritme Rocchio. Terdapat aturan dalam penentuan kelas sebagai berikut:

1. kelas positif = positif >= negatif && positif >= netral
2. kelas netral = netral >= negatif && netral >= positif
3. kelas negatif = netral && negatif >= positif

TABEL 5  
KLASIFIKASI *SENTIMENT* ANALISIS

kata	Netral	Negatif	Positif	kelas
ada	0.05	0.091	0.100	positif
adalah	0.075	0.037	0.050	netral
adanya	0	0.003	0.016	positif
adobe	0.125	0.071	0.051	netral
aduh	0	0.027	0	negatif
agak	0.025	0.025	0.041	negatif
agar	0	0.021	0.006	negatif

Tabel 5 menunjukkan daftar kata dengan masing-masing memiliki nilai *centroid* (tengah). Nilai *centroid* tersebut dibandingkan dengan nilai *centroid* netral, nilai *centroid* negatif dan nilai *centroid* positif. Jika nilainya lebih besar dari nilai *centroid* lainnya maka kata tersebut merupakan kelas yang diambil. Model rekomendasi dilakukan dengan cara mempertimbangkan opini *sentiment* twitter dengan nilai tugas provek.

### E. Kesimpulan

Dalam penelitian ini telah dibuat sebuah model rekomendasi dengan menggunakan *opinion analytics* untuk evaluasi *project based learning* mata kuliah teknik multimedia. Hasil penelitian menunjukkan bahwa model sistem mampu memberikan rekomendasi yang diharapkan dapat meningkatkan kinerja dosen dan meningkatkan pelaksanaan perkuliahan mata kuliah teknik multimedia.

#### REFERENSI

- [1] Pang B, Lee L. *Opinion Mining and Sentimen Analysis. Foundations and Trends R in Information Retrieval*. 2(1–2): hlm 1–13. 2008
- [2] Liu, Bing. *Sentiment Analysis and Opinion Mining Chicago (US): University of Illinois*. 2012
- [3] Novantirani A. Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat dalam Kota dengan Metode Support Vector Machine. *Journal e-Proceeding of Engineering*. 2(1): 1177. 2015
- [4] Nurzahputra A, Muslim MA. Analisis Sentiment Pada Opini Mahasiswa Menggunakan Natural Language Processing. *Seminar Nasional Ilmu Komputer*. Semarang (ID): hlm 114-118. 2016
- [5] Pan G, Seow PS, Koh G. *Xamining Learning Transformativ in ProjectBased Learning Process*. *Journal of International Education in Business*. 12, (2), 167-180. Research Collection Shool of Accountancy. 2019
- [6] Misuraca M, Spano M, Scepi G. *Using Opinion Mining as an educational analytic: An integrated strategy for the analysis of students' feedback*. Artikel di *Studies In Educational Evaluation*. doi: 10.1016/j.stueduc.2021.100979. 2021
- [7] Pantouw JC. Perbandingan Klasifikasi Rocchio dan Multinomial Naive Bayes pada Analisis Sentimen Data Twitter Bahasa Indonesia. [skripsi]. Bogor (ID): Departemen Ilmu Komputer, Institut Pertanian Bogor. hlm 36. 2017
- [8] Pantouw JC. Perbandingan Klasifikasi Rocchio dan Multinomial Naive Bayes pada Analisis Sentimen Data Twitter Bahasa Indonesia. [skripsi]. Bogor (ID): Departemen Ilmu Komputer, Institut Pertanian Bogor. hlm 36. 2017