

## **Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes**

**Yuni Nurtikasari<sup>1</sup>, Syariful Alam<sup>2\*</sup>, Teguh Iman Hermanto<sup>3</sup>**

<sup>1,2\*,3</sup>Teknik Informatika, Sekolah Tinggi Teknologi Wastukancana, Purwakarta, Indonesia

Email: <sup>1</sup>yuninurtikasari21@wastukancana.ac.id, <sup>2\*</sup>syarifulalam@wastukancana.ac.id,

<sup>3</sup>teguhiman@wastukancana.ac.id

### **Abstract**

*Film is one of the most interesting topics to talk about. When someone writes an opinion on a film, all the elements in the film will be written down. Film opinion data in this study were taken from film comments written on twitter. The number of opinions written on Twitter requires classification according to the sentiments they have so that it is easy to get the tendency of the opinion towards the film whether it tends to have a positive, negative or neutral opinion. Recently, the spotlight on twitter media in Indonesia is a film with the title Horror-Ngeri Sedap. Ngeri-Ngeri Sedap is a 2022 Indonesian comedy-drama film directed and written by Bene Dion Rajagukguk. The film is set in the Batak Tribe, starring Arswendy Beningswara Nasution, Tika Panggabean, Boris Bokir Manullang, Gita Bhebhita Butar-butar, Lolox, and Indra Jegel. This causes differences in the views and opinions of twitter users towards the Horror Sedap film. So it is necessary to have a sentiment classification for the opinion. The use of the Naïve Bayes Algorithm was chosen in the analysis because it has the highest probability or opportunity value for data classification. Labeling on Twitter data is done manually by giving positive, negative, neutral sentiments to the raw dataset in Microsoft excel then the data enters the preprocessing transformation, tokenization and filtering stages. The tf-idf weighting is carried out when the data is complete in the transformation process, tf-idf is used to determine the number of occurrences of words, then the data classification is carried out using the Naïve Bayes Algorithm. confusion matrix testing is done after the data classification is complete using the orange tools. Based on the test results with the confusion matrix with orange tools, the average accuracy value is 0.65% and the precision value is 0.67%, and recall is 0.65%, and the percentage is neutral 0.83% in its classification. This proves that the public sentiment on the twitter platform towards the case of the film Horror Sedap is neutral and the Naïve Bayes Algorithm is considered reliable and valid in data processing.*

**Keywords:** Film, Naïve Bayes, Sentiment Analysis, Orange.

### **Abstrak**

Film merupakan salah satu topik yang sangat menarik untuk dibicarakan. Ketika seseorang menulis opini suatu film, maka semua unsur yang ada di dalam film tersebut akan dituliskan. Data opini film pada penelitian ini diambil dari komentar film yang ditulis di twitter. Banyaknya opini yang dituliskan di twitter membutuhkan pengklasifikasian sesuai sentimen yang dimiliki agar mudah untuk mendapatkan kecenderungan opini tersebut terhadap film apakah cenderung beropini positif, negatif atau netral. Akhir-akhir ini jadi sorotan di media twitter di Indonesia adalah film dengan judul Ngeri-Ngeri Sedap. Ngeri-Ngeri Sedap adalah film drama komedi Indonesia tahun 2022 yang disutradarai dan ditulis oleh Bene Dion Rajagukguk. Film berlatar Suku Batak ini dibintangi oleh Arswendy Beningswara Nasution, Tika Panggabean, Boris Bokir Manullang, Gita Bhebhita Butar-butar, Lolox, dan Indra Jegel. Hal tersebut menyebabkan adanya perbedaan pandangan dan pendapat dari pengguna twitter terhadap film Ngeri-ngeri Sedap. sehingga diperlukan adanya klasifikasi sentimen untuk opini tersebut. Penggunaan Algoritma Naïve Bayes dipilih dalam analisis karena mempunyai nilai probabilitas atau peluang tertinggi untuk

pengklasifikasian data. Pelabelan pada Data *twitter* dilakukan secara manual dengan pemberian sentimen positif, negatif, netral pada *dataset* mentah di *Microsoft excel* kemudian data masuk dalam tahap *preprocessing transformation, tokenization* dan *filtering*. Pembobotan tf-idf dilakukan ketika data selesai pada proses transformasi, tf-idf digunakan untuk mengetahui banyaknya kemunculan kata, kemudian pengklasifikasian data dilakukan dengan Algoritma *Naïve Bayes*. pengujian *confusion matrix* dilakukan setelah pengklasifikasian data selesai menggunakan *tools orange*. Berdasarkan Hasil pengujian dengan *confusion matrix* dengan *tools orange* didapatkan hasil rata – rata nilai akurasi 0.65% dan nilai presisi sebesar 0.67%, dan *recall* sebesar 0.65%, dan persentase netral 0.83% pada klasifikasi nya. Hal ini membuktikan bahwa sentimen masyarakat pada *platform twitter* terhadap kasus film Ngeri-ngeri Sedap tergolong netral dan Algoritma *Naïve Bayes* dinilai reliabel dan valid dalam pemrosesan data.

**Kata Kunci:** Film, *Naïve Bayes*, Analisis Sentimen, *Orange*.

## 1. PENDAHULUAN

Perkembangan dunia teknologi informasi dan komunikasi yang pesat tidak terlepas dari penyedia layanan *website*, kemudahan manusia dalam menggunakan *website* mengakibatkan bertambahnya dokumen teks yang berupa pendapat dan informasi yang beragam. Dalam waktu yang lama dokumen teks akan bertambah besar. Banyaknya dokumen teks berasal dari berbagai sumber seperti *review*, opini, berita, *paper*, buku perpustakaan digital, pesan *e-mail* dan halaman *website*.

Saat ini menonton film sudah menjadi hal yang dilakukan dalam mengisi waktu luang. Semakin banyak film yang bermunculan baik dalam maupun luar negeri membuat penonton film memiliki banyak pilihan dan mencari informasi dengan membaca opini tentang film tersebut. Pendapat orang-orang dapat mengurangi ketidakpastian terhadap suatu film tertentu dan membantu penonton menyimpulkan kualitas suatu film tertentu. Saat ini, pendapat khalayak umum menjadi sumber yang penting dalam pengambilan keputusan seseorang akan suatu film. Tapi terkadang selera membuat faktor yang lainnya tidak masuk hitungan.

Film adalah suatu bentuk komunikasi massa elektronik yang berupa media audio visual yang mampu menampilkan kata-kata, bunyi, citra, dan kombinasinya. Film juga merupakan salah satu bentuk komunikasi modern yang kedua muncul di dunia. Film berperan sebagai sebuah sarana baru yang digunakan untuk menyebarkan hiburan yang sudah menjadi kebiasaan terdahulu, serta menyajikan cerita, peristiwa, musik, drama, lawak, dan sajian teknis lainnya kepada masyarakat umum. Film juga menurut Prof.Effendy adalah media komunikasi massa yang ampuh sekali, bukan saja untuk hiburan, tetapi juga untuk penerangan dan pendidikan. Film mempunyai suatu dampak tertentu terhadap penonton, dampak – dampak tersebut dapat berbagai macam seperti, dampak psikologis, dan dampak sosial (Oktavianus, 2015).

Beberapa opini tentang film dapat membantu penonton dalam mengetahui kualitas film apakah sangat dianjurkan, direkomendasikan dan tidak terlalu dianjurkan untuk ditonton. Saat ini penonton yang menulis opini dan pengalaman secara online semakin meningkat. Opini film dapat memberikan manfaat yang besar bagi penonton karena dengan membaca opini, penonton mendapatkan informasi yaitu dapat mengetahui kualitas suatu film dari opini penonton lain yang membagikan pengalamannya tentang film terkait.

Industri film secara global terus mengalami perkembangan, baik dari jumlah film yang dihasilkan, jumlah penonton, maupun jumlah perputaran uangnya. Data yang didapatkan dari *National Association of Theater Owners* (asosiasi pemilik bioskop di Amerika Serikat) menunjukan bahwa ada perkembangan dari tahun 1987 dimana tiket

bioskop terjual sebanyak 1,09 milliar tiket, menjadi 1,314 milliar tiket pada tahun 2016 untuk penjualan di wilayah Amerika Serikat dan Kanada. Sedangkan untuk pendapatan *box office* di wilayah Amerika Serikat dan Kanada pada tahun 1987 sebesar \$4,25 miliar menjadi \$11,372 miliar pada tahun 2016, angka ini menunjukan peningkatan yang sangat signifikan (Mandiri, 2018).

Informasi yang menyebabkan pertambahan data yang kebanyakan berupa data teks dapat dijadikan sumber yang sangat potensial untuk digali lebih dalam. Teknik yang berkembang untuk penggalian dokumen teks saat ini adalah *text mining*. Oleh karena itu proses analisis sentimen menjadi sangat penting agar dapat mengetahui seberapa akuratkah suatu opini film tersebut.

*Texts mining* merupakan proses ekstraksi pola (informasi dan pengetahuan yang berguna) dari sejumlah data tak terstruktur yang nantinya akan diperoleh pola-pola data, tren dan ekstraksi pengetahuan yang potensial dari data teks. Masukan untuk penambangan teks adalah data yang tidak (atau kurang) terstruktur, seperti dokumen, *word*, PDF, kutipan teks dll sedangkan masukan untuk penambangan data adalah data yang terstruktur. Salah satu tujuan penggunaan *text mining* adalah analisis sentimen (Ratnawati, 2018)

Ada beberapa kategori yang termasuk didalam teknik *text mining* salah satunya adalah analisis sentimen, yaitu suatu proses memahami, mengekstrak, dan mengolah data textual secara otomatis, atau merupakan studi komputasi pendapat, perasaan dan emosi yang dinyatakan dalam bentuk teks. Analisis sentimen berguna untuk menentukan apakah opini atau komentar terhadap permasalahan memiliki kecenderungan positif, negatif atau netral dan dapat dijadikan sebagai acuan dalam meningkatkan suatu pelayanan, ataupun meningkatkan kualitas tertentu. Besarnya manfaat serta pengaruh dari melakukan analisis sentimen membuat penelitian dan aplikasi mengenai analisis sentimen berkembang pesat. Oleh karena itu, analisis sentimen merupakan salah satu solusi mengatasi masalah untuk mengelompokan opini atau opini menjadi opini positif, negatif dan netral.

Analisis sentimen adalah studi yang bertujuan untuk menganalisis opini, sentimen dan emosi yang terdapat pada dokumen atau data. Tugas dasar dari analisis sentimen adalah untuk mengelompokkan sifat dari teks yang ada di dalam kalimat maupun pendapat yang pada umumnya terbagi menjadi kelas yaitu positif dan negatif. Besarnya pengaruh dan manfaat dari analisis sentimen menyebabkan penelitian dan aplikasi berbasis analisis sentimen berkembang pesat. Bahkan di Amerika terdapat sekitar 20-30 perusahaan yang memfokuskan pada layanan analisis sentimen (Afrizal, 2019).

Penelitian ini membahas bagaimana cara memanfaatkan opini pengguna *Twitter* terhadap film untuk analisis sentimen, kemudian membagi opini tersebut menjadi data sentimen positif, negatif dan netral. Pengklasifikasian kumpulan data *tweet* tersebut bisa menggunakan salah satu metode klasifikasi yaitu *Naïve Bayes Classifier*. Nantinya, bisa digunakan untuk memberi informasi tentang film mana yang direkomendasikan agar mereka tonton. Terkadang ada film yang disukai penonton tapi tidak disukai kritikus, terkadang para kritikus suka tapi penonton tidak suka dan terkadang baik penonton maupun kritikus menyukai filmnya.

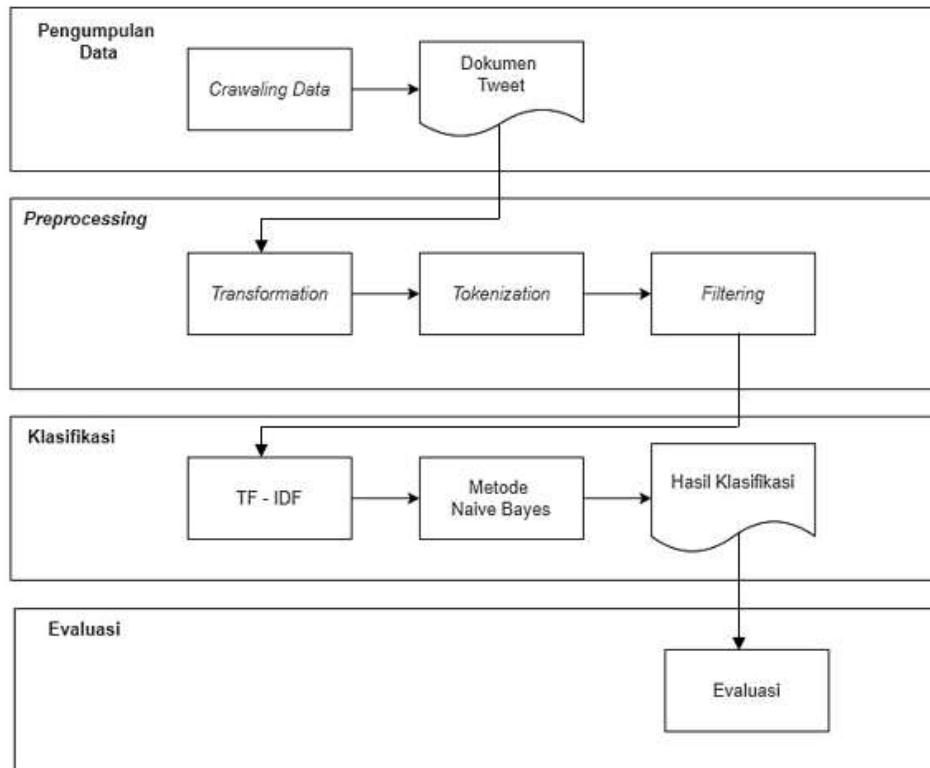
Sedangkan untuk aplikasi pengolah data, *tools orange* dapat digunakan dikarenakan sudah mendukung berbagai algoritma dan metode di dalamnya.

## 2. METODOLOGI PENELITIAN

### 2.1 Kerangka Penelitian

Kerangka pikir merupakan jalur pemikiran yang dirancang berdasarkan kegiatan peneliti yang dilakukan. Kerangka berpikir merupakan model konseptual tentang bagaimana teori berhubungan dengan berbagai faktor yang telah diidentifikasi sebagai masalah yang penting (Sugiyono, 2018).

Rancangan penelitian yang akan dilakukan disajikan dalam bentuk kerangka pemikiran sebagai berikut :



Gambar 2.1 Kerangka Penelitian

1. Pada tahap *crawling data* dilakukan penarikan data *text Twitter* untuk menarik data (*crawling*) dengan *mention Negeri-negeri Sedap*.
2. Pada tahap *labelling* dilakukan pemberian label positif, negatif dan netral pada *text* dokumen.
3. Pada tahap *preprocessing* dilakukan *transformation*, *tokenizing*, *filtering* menyiapkan dokumen mentah menjadi dokumen atau representatif dokumen yang siap diproses.
4. Pada tahap TF-IDF dilakukan pembobotan pada kata dalam dokumen untuk mengetahui bobot dari kata tersebut.
5. Pada tahap *Classification* dilakukan pengklasifikasian data dengan menggunakan algoritma *naïve bayes*.
6. Pada tahap *Evaluation* dilakukan perhitungan *accuracy*, *precision*, *recall* dengan *confusion matrix* pada dokumen yang telah terkласifikasi algoritma *naïve bayes*.

## 2.2 Objek Penelitian

Pada penelitian ini, dataset yang diambil berasal dari *twitter* dengan objek yang diambil yaitu 1 judul film yaitu Ngeri-ngeri Sedap. Dengan subjek yaitu data pengguna *twitter*. Data ini diambil menggunakan *tools Orange*.

## 2.3 Teknik Pengumpulan Data

Proses pengumpulan data dalam penelitian ini mengenai review film yang di peroleh dari *twitter* dengan *query* “Ngeringeri Sedap” dengan metode *crawling data*. Data *tweet* yang diambil merupakan *tweet* dari akun pengguna *twitter*.

## 2.4. Preprocessing Text

*Preprocessing Text* membagi teks sehingga menjadi unit-unit yang lebih kecil seperti *transformation*, *tokenization*, *filtering*. Langkah – langkah dalam analisis adalah diterapkan secara berurutan dan dapat diaktifkan atau dinonaktifkan. *Preprocess text* akan melaporkan beberapa hal seperti jumlah dokumen (tentang jumlah dokumen yang di masukkan), *total token* (menghitung semua *token* dalam *corpus* (kumpulan dokumen), dan *token* yang dilaporkan hanya pada *token* unik di *corpus* bukan *token* duplikat.

## 2.5 Pembobotan Kata (TF-IDF)

Setelah *preprocessing* (*transformation*, *tokenization*, *filtering*) pada dokumen selesai, selanjutnya dilakukan proses *term weighting*, yang akan diberi bobot (nilai), dimana bobot tersebut menunjukkan penting nya *term* terhadap dokumen. Selanjutnya dilakukan perhitungan bobot pada *term* yang dicari pada setiap dokumen bertujuan untuk mengetahui kegunaan dan kesamaan dari *term* dalam dokumen. Semakin banyak kata yang muncul dalam kumpulan dokumen, semakin tinggi nilai atau bobot kata tersebut. Setelah tahap pembobotan selesai dilakukan proses klasifikasi.

# 3. HASIL DAN PEMBAHASAN

## 3.1. Klasifikasi Naïve Bayes

Dokumen *tweet* yang telah dilakukan tahap *preprocessing* dapat diklasifikasikan ke dalam kategori tertentu berdasarkan kata – kata yang terkandung dalam dokumen tersebut. Kumpulan kata pada dokumen dapat digunakan untuk menentukan kategorinya karena kata – kata tersebut dinilai memiliki makna tertentu.

### Menghitung Klasifikasi

Berikut ini adalah contoh kumpulan data yang akan digunakan hingga 9 data *tweet*, termasuk 9 sampel yang ditandai sebagai 3 positif, 3 negatif, dan 3 netral. Data tersebut digunakan sebagai data latih, dapat dilihat pada tabel dibawah ini :

Tabel 3.1 Data Sampel

Dokumen	Teks	sentimen
D1	film indonesia dengan sinematografi terbaik	positif
D2	film drama keluarga chemistry kuat	positif
D3	sukses film ngeri ngeri sedap mengangkat budaya batak	positif
D4	awas kalo gita bebita dapat penghargaan	negatif
D5	film roller coaster	negatif
D6	ngeri ngeri sedap akan pendidikan pahit	negatif
D7	makin tua makin ambyar nonton film tema keluarga	netral
D8	lucu dapet banget sedih dapet banget	netral
D9	jadi iri keluarga bapak domu luar biasa keren	netral

Dari data tabel 3.1 sampel dataset diatas dilakukan perhitungan dengan menggunakan *Microsoft Excel* yang pertama dilakukan adalah menentukan *term frequency (TF)* pada setiap dokumen yang dapat dilihat pada tabel 3.2 dibawah ini.

Tabel 3. 1 *Term Frequency*

Kata	D1	D2	D3	D4	D5	D6	D7	D8	D9
film	1	1	1	0	1	0	1	0	0
indonesia	1	0	0	0	0	0	0	0	0
dengan	1	0	0	0	0	0	0	0	0
sinematografi	1	0	0	0	0	0	0	0	0
terbaik	1	0	0	0	0	0	0	0	0
drama	0	1	0	0	0	0	0	0	0
keluarga	0	1	0	0	0	0	1	0	1
chemistry	0	1	0	0	0	0	0	0	0
kuat	0	1	0	0	0	0	0	0	0
sukses	0	0	1	0	0	0	0	0	0
ngeri	0	0	2	0	0	2	0	0	0
sedap	0	0	1	0	0	1	0	0	0
mengangkat	0	0	1	0	0	0	0	0	0
budaya	0	0	1	0	0	0	0	0	0
batak	0	0	1	0	0	0	0	0	0
awas	0	0	0	1	0	0	0	0	0
kalo	0	0	0	1	0	0	0	0	0
gita	0	0	0	1	0	0	0	0	0
bebita	0	0	0	1	0	0	0	0	0
dapat	0	0	0	1	0	0	0	0	0
penghargaan	0	0	0	0	0	0	0	0	0
roller	0	0	0	0	1	0	0	0	0
coaster	0	0	0	0	1	0	0	0	0
akar	0	0	0	0	0	1	0	0	0
pendidikan	0	0	0	0	0	1	0	0	0
pahit	0	0	0	0	0	1	0	0	0

makin	0	0	0	0	0	0	1	0	0
tua	0	0	0	0	0	0	1	0	0
ambayar	0	0	0	0	0	0	1	0	0
nonton	0	0	0	0	0	0	1	0	0
tema	0	0	0	0	0	0	1	0	0
lucu	0	0	0	0	0	0	0	1	0
dapet	0	0	0	0	0	0	0	2	0
banget	0	0	0	0	0	0	0	2	0
sedih	0	0	0	0	0	0	0	1	0
jadi	0	0	0	0	0	0	0	0	1
iri	0	0	0	0	0	0	0	0	1
bapak	0	0	0	0	0	0	0	0	1
domu	0	0	0	0	0	0	0	0	1
luar	0	0	0	0	0	0	0	0	1
biasa	0	0	0	0	0	0	0	0	1
keren	0	0	0	0	0	0	0	0	1

selanjutnya dilakukan perhitungan jumlah dari kemunculan kata unik dan jumlah kata pada setiap kategori. Alur perhitungan dapat dilihat pada tabel 3.3 dibawah ini :

Tabel 3.3 Kata Unik

Kategori	Jumlah dokumen di setiap kategori	Jumlah dokumen seluruhnya	P(Kategori)	Jumlah Kata Setiap Kategori
positif	3	9	0.33	17
negatif	3	9	0.33	14
Netral	3	9	0.33	19

Jumlah Kata Unik	50
------------------	----

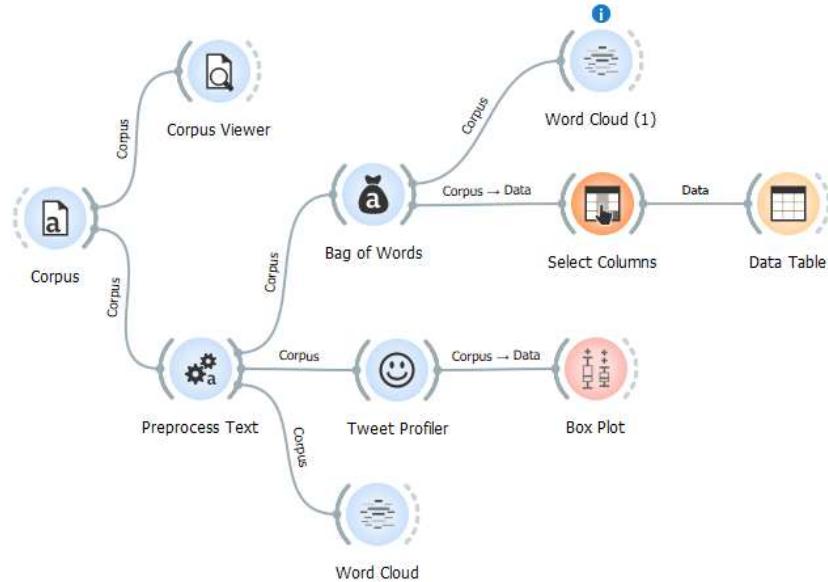
Kata unik merupakan kata yang sering muncul di setiap dokumen. Setelah mendapatkan jumlah kata unik dan jumlah kata pada setiap kategori, yang dilakukan selanjutnya adalah menghitung probabilitas pada setiap kategori.

$$P(\text{Kata}) = \frac{\text{jumlah dari satu kata di satu kategori} + 1}{\text{jumlah kata tiap kategori} + \text{jumlah keseluruhan kata unik}}$$

Dari tabel *term frequency* (TF) di atas didapatkan hasil jumlah kata, dalam kategori positif 17 kata, negatif 14 kata dan netral 19 kata dengan jumlah kata unik sebanyak 50 kata.

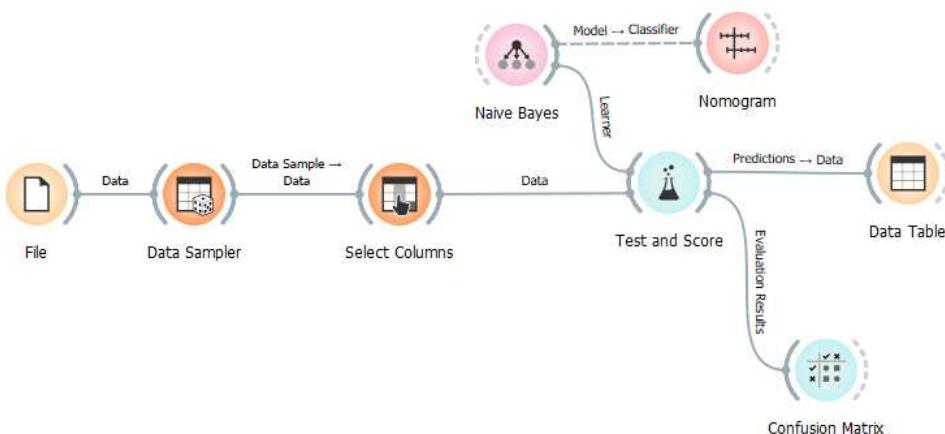
## Pengujian Naïve Bayes Menggunakan Orange

Berikut merupakan alur pengujian *Naïve Bayes* menggunakan tools *orange* yang dapat dilihat pada gambar



Gambar 3.1 Alur pengujian Naïve Bayes menggunakan tools orange

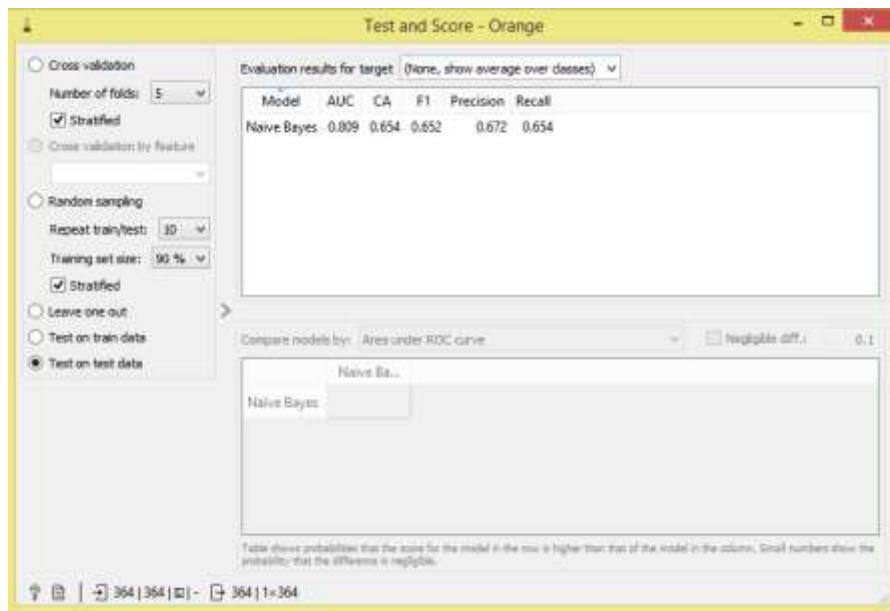
Berikut Merupakan alur pengujian data sampler menggunakan tools *orange* yang dapat dilihat pada gambar



Gambar 3.2 Alur pengujian data sampler menggunakan tools orange

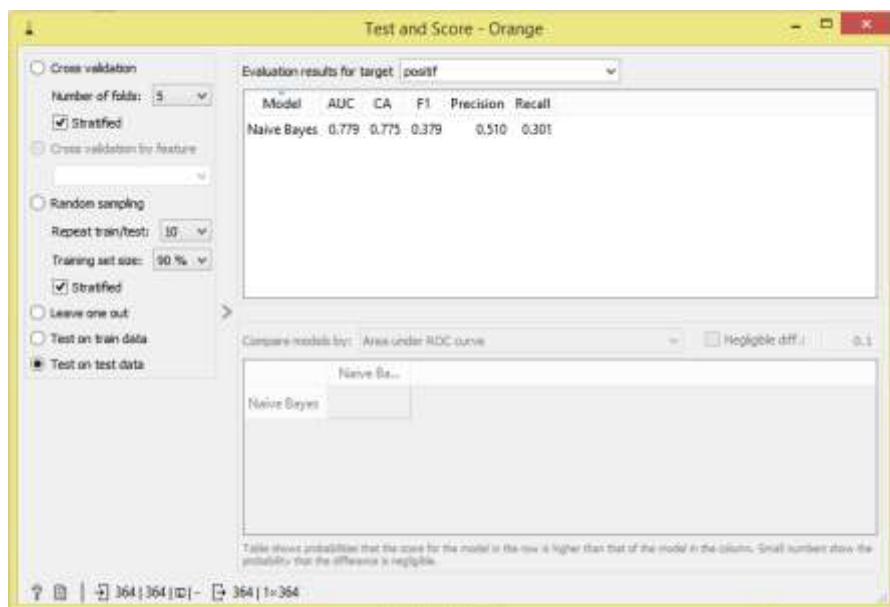
## Hasil Klasifikasi Algoritma Naïve Bayes

Perhitungan *accuracy*, *precision*, dan *recall* dengan sistematisasi digunakan untuk membuktikan hasil dari performa algoritma *naïve bayes* dengan menggunakan *tool orange*. Hasil klasifikasi dapat dilihat pada gambar 3.3 sampai 3.6 dibawah ini.



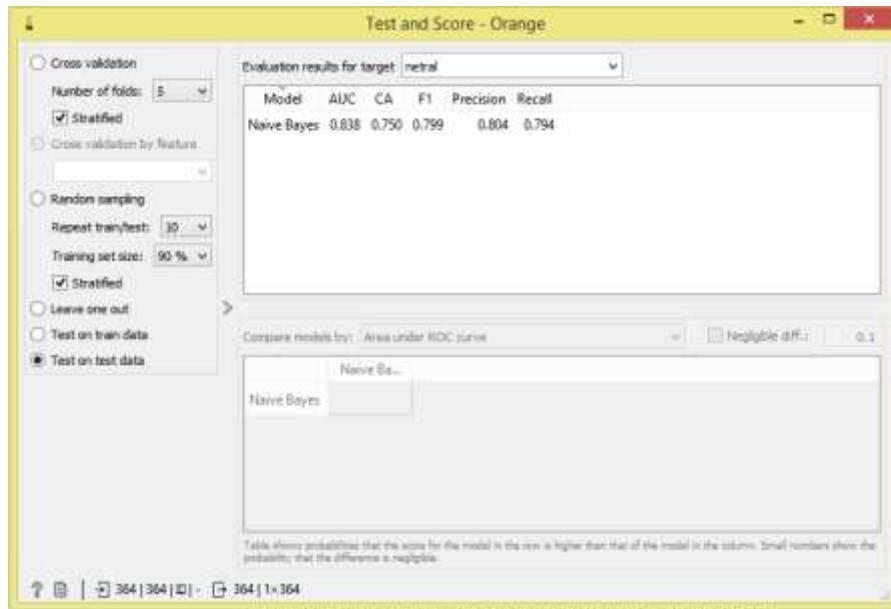
Gambar 3. 3 Klasifikasi Rata-Rata Dari Tiap Kelas Sentimen

Pada tabel diatas menunjukkan hasil dari perhitungan *accuracy*, *precision*, dan *recall* pada rata-rata tiap kelas sentimen.



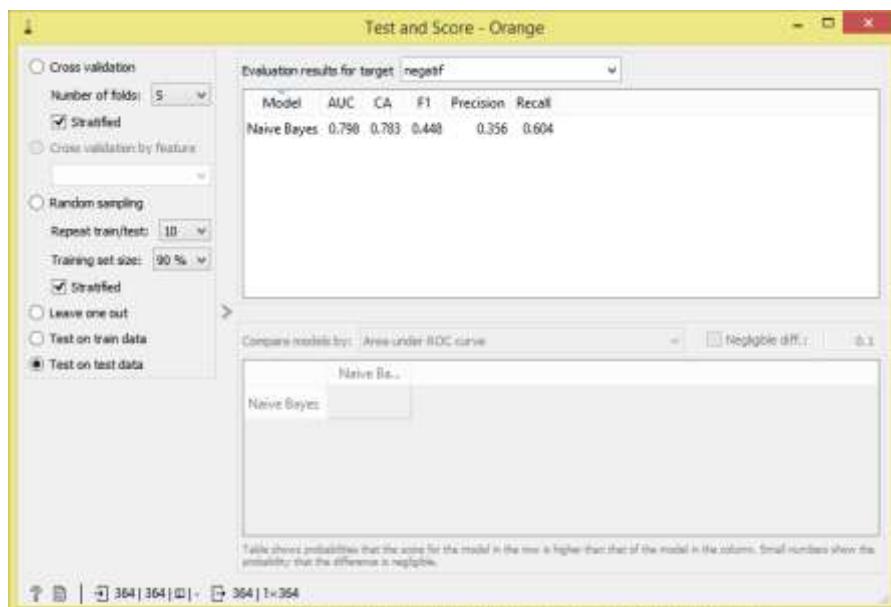
Gambar 3.4 Klasifikasi Rata-Rata Dari Tiap Kelas Sentimen

Pada tabel diatas menunjukkan hasil dari perhitungan *accuracy*, *precision*, dan *recall* pada rata-rata tiap kelas sentimen.



Gambar 3. 5 Klasifikasi Rata-Rata Dari Tiap Kelas Sentimen

Pada tabel diatas menunjukkan hasil dari perhitungan *accuracy*, *precision*, dan *recall* pada rata-rata tiap kelas sentimen.



Gambar 3. 6 Klasifikasi Rata-Rata Dari Tiap Kelas Sentimen

Pada tabel diatas menunjukkan hasil dari perhitungan *accuracy*, *precision*, dan *recall* pada rata-rata tiap kelas sentimen.

### 3.2. Evaluasi

Evaluasi bertujuan untuk mengecek kebenaran hasil klasifikasi. Pada tahap ini hasil evaluasi menggunakan metode *cross validation* yang nantinya akan diukur dengan menggunakan *Confusion Matrix* untuk mengetahui hasil akurasi, presisi, dan *recall*.

Berikut merupakan tabel dari *confusion matrix* hasil yang di dapatkan setelah dilakukan pengujian dari keseluruhan data tweet dengan menggunakan *tools orange* dapat dilihat pada tabel 3.3

Tabel 3.3 confusion matrix

		PREDIKSI			
		NEGATIF	NETRAL	POSITIF	
AKTUAL	NEGATIF	32	15	6	53
	NETRAL	29	181	18	228
	POSITIF	29	29	25	83
		90	225	49	364

Perhitungan *Akurasi*

$$A = \frac{TP + TN_{netral} + TN}{\text{Jumlah Data}} = \frac{25 + 181 + 32}{364} = \frac{238}{364} = 0,65$$

Perhitungan *Precision*

Tabel 3.4 Perhitungan *Precision*

	Negatif	Netral	Positif
TP	32	181	18
FP	15 + 6	29 + 18	29 + 29
Precision	32 / (32 + 21)	181 / (181 + 47)	18 / (18 + 58)
	0.60	0.79	0.23

$$All\ Precision = \frac{Precision\ A + B + C}{\text{Jumlah Kelas}}$$

$$All\ Precision = \frac{0.60 + 0.79 + 0.23}{3} = \frac{1.53}{3} = 0.51$$

Perhitungan *Recal*

Tabel 3.5 Perhitungan *Recal*

	Negative	Netral	Positif
TP	32	181	18
FP	29 + 29	15 + 29	6 + 18
Recall	32 / (32 + 58)	181 / (181 + 44)	18 / (18 + 24)
	0.33	0.80	0.42

$$Recall = \frac{Recall\ A + B + C}{\text{Jumlah Kelas}}$$

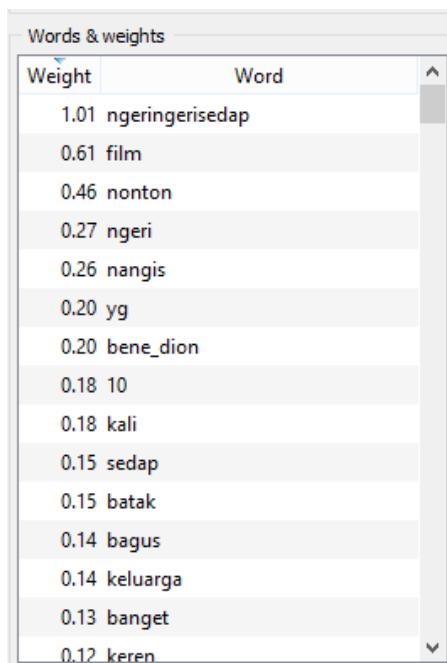
$$Recall = \frac{0.33 + 0.80 + 0.42}{3} = \frac{1.55}{3} = 0.51$$

### 3.3 Visualisasi

Visualisasi adalah suatu rekayasa dalam pembuatan gambar, diagram atau animasi untuk penampilan suatu informasi. Visualisasi digambarkan dengan *word cloud* pada sentimen keseluruhan dokumen negatif positif dan netral tentang film Ngeri-ngeri Sedap. Hasil dari Visualisasi ditampilkan pada gambar dibawah ini :



Gambar 3.3 Hasil Word Cloud



Gambar 3.4 Hasil Words & Weights

#### 4. KESIMPULAN

Dari hasil penelitian diatas mengenai analisis sentimen mengenai film Ngeri-ngeri Sedap pada *platform twitter* yang berjumlah 404 data dan sudah melewati tahap *preprocess text* yang terdiri dari *filtering*, *tokenization*, *transformation* dan *klasifikasi* menggunakan algoritma *naïve bayes* serta evaluasi data dengan *confusion matrix* menggunakan *tools Orange* dapat disimpulkan bahwa opini masyarakat mengenai film Ngeri-ngeri Sedap tergolong netral dengan persentase 75% pada akurasi (*accuracy*),

80% pada nilai *precision*, dan tingkat keberhasilan (*recall*) 79%. Berdasarkan nilai tersebut membuktikan sentimen masyarakat mengenai film Ngeri-ngeri Sedap pada *platform twitter* tergolong netral.

## REFERENCES

- (1). Afrizal, Sarika., Irmanda, Helena. N., Falih, N., & Isnainiyah, I. N. (2019). Implementasi Metode *Naïve Bayes* untuk Analisis Sentimen Warga Jakarta Terhadap Kehadiran *Mass Rapid Transit*. *Jurnal Informatik*, 4221, 157–168.
- (2). Mandiri, S. N. (2018). *KOMPARASI ALGORITMA KLASIFIKASI UNTUK ANALISIS SENTIMEN REVIEW FILM BERBAHASA ASING*. 180–185.
- (3). Oktavianus, H. (2015). Penerimaan Penonton Terhadap Praktek Eksorsis Di Dalam Film Conjuring. *E-Komunikasi*, 3(2), 12. <https://media.neliti.com/media/publications/79600-ID-none.pdf>
- (4). Ratnawati, F. (2018). *Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter*.
- (5). Sugiyono, P. D. (2018). *METODE PENELITIAN KUANTITATIF KUALITATIF DAN R&D*.