

# PENGEMBANGAN SISTEM ANALISA KEBERPIHAKAN MEDIA ONLINE BERDASARKAN TREND WAKTU MENGGUNAKAN NAIVE BAYES CLASSIFIER

Faisal Rahutomo<sup>1</sup>, Annisa Taufika Firdausi<sup>2</sup>, Nur Rochmanshah<sup>3</sup>

<sup>1,2,3</sup> Teknik Informatika, Teknologi Informasi, Politeknik Negeri Malang

<sup>1</sup>faisal@polinema.ac.id, <sup>2</sup>annisa.taufika@polinema.ac.id, <sup>3</sup>rochmanshah100@gmail.com

---

## Abstrak

Media *online* menjadi salah satu media yang paling penting dalam menyajikan informasi kepada para pembaca. Media *online* dapat menyajikan informasi kepada pembaca dengan cepat, gratis, dan praktis. Salah satu topik yang menjadi perhatian dari masyarakat Indonesia, khususnya di tahun 2019 adalah topik yang berkaitan dengan pemilihan Presiden dan Wakil Presiden. Banyak media *online* dalam penyajian berita di setiap hari, secara terang-terangan maupun tidak, berpihak kepada salah satu Pasangan Presiden dan Wakil Presiden. Penyajian Media *Online* yang tidak netral ataupun tidak objektif tidak hanya merugikan bagi pihak Pasangan Presiden dan Wakil Presiden, tetapi juga mampu memberikan perspektif berbeda bagi pembaca atau masyarakat kepada pasangan Presiden dan Wakil Presiden terkait. Salah satu cara untuk menentukan atau mengetahui keberpihakan media kepada pasangan Presiden dan Wakil Presiden adalah dengan membuat sistem yang secara otomatis dapat mengklasifikasikan berita-berita yang ada pada suatu media di setiap hari atau waktu tertentu menjadi kategori yang telah ditentukan. Hasil pengklasifikasian kemudian diolah dan disajikan dalam bentuk grafik serta dihitung nilai keberpihakannya. Salah satu metode pengklasifikasian adalah dengan menggunakan metode *Naive Bayes Classifier* (NBC). NBC adalah klasifikasi statistik yang bisa memprediksi probabilitas sebuah kelas, dan kelebihan dari metode ini adalah tingkat akurasi yang tinggi juga waktu komputasi yang lebih cepat.

**Kata Kunci :** Analisa Keberpihakan Media *Online*, Pemilihan Presiden dan Wakil Presiden, *Naive Bayes Classifier*.

---

## 1. Pendahuluan

Media *online* adalah sebutan umum untuk sebuah bentuk media yang berbasis telekomunikasi dan multimedia (baca-komputer dan internet). Didalamnya terdapat portal, *website* (situs *web*), *radio-online*, *TV-online*, *pers online*, *mail-online*, dll, dengan karakteristik masing-masing sesuai dengan fasilitas yang memungkinkan user memanfaatkannya. Media *online* adalah media massa yang tersaji secara *online* di situs *web* (*website*) internet. Media *online* adalah media massa "generasi ketiga" setelah media cetak (*printed media*) koran, tabloid, majalah, buku, dan media elektronik (*electronic media*) radio, televisi, dan film atau video.

Di zaman modern, media *online* menjadi salah satu media yang paling penting dalam menyajikan informasi kepada para pembaca, Media *online* dapat menyajikan informasi kepada pembaca dengan cepat, gratis, dan praktis, bahkan untuk mengakses media *online* pun dapat melalui *smartphone* ataupun komputer yang terhubung dengan internet. Hal ini tentu memberikan manfaat kepada para pembaca dalam mendapatkan informasi terbaru di setiap harinya.

Dalam menyajikan berita, media *online* menyajikan berita-berita sesuai dengan beragam topik yang diminati oleh masyarakat, mulai dari yang berkaitan dengan Ekonomi, Olahraga, Bisnis, Teknologi, Kesehatan, Politik, dan sebagainya. Salah satu topik yang menjadi perhatian dari masyarakat Indonesia, khususnya di tahun 2019 adalah topik mengenai politik, terutama yang berkaitan dengan pemilihan Presiden dan Wakil Presiden. Hal tersebut juga dikarenakan bertepatan dengan diselenggarakannya pemilihan Presiden dan Wakil Presiden di tahun 2019.

Dalam hal yang berkaitan dengan Pemilihan Presiden dan Wakil Presiden 2019. Media *online* dapat berperan sebagai media yang menyampaikan informasi yang berkaitan dengan penyelenggaraan Pemilihan Presiden dan Wakil Presiden 2019, ataupun yang berkaitan dengan hal-hal politik maupun non politik yang dilakukan oleh masing-masing pasangan Presiden dan Wakil Presiden. Pada hakikatnya, media itu memang harus objektif dalam melihat isu dan juga dalam pemberitaannya. Akan tetapi, pada kenyataannya justru banyak pemberitaan mengenai pasangan Presiden dan Wakil Presiden yang justru tidak netral atau tidak objektif, memihak

kepada salah satu pasangan, bahkan menyudutkan pasangan yang lain. Banyak media *online* dalam penyajian berita di setiap hari, secara terang-terangan maupun tidak, berpihak kepada salah satu Pasangan Presiden dan Wakil Presiden. Penyajian Media *Online* yang tidak netral ataupun tidak objektif tidak hanya merugikan bagi pihak Pasangan Presiden dan Wakil Presiden, tetapi juga mampu memberikan perspektif berbeda bagi pembaca atau masyarakat kepada pasangan Presiden dan Wakil Presiden terkait. Baik itu berupa perspektif positif ataupun perspektif negatif.

Salah satu cara untuk menentukan atau mengetahui keberpihakan media kepada pasangan Presiden dan Wakil Presiden adalah dengan membuat sistem yang secara otomatis dapat mengklasifikasikan berita-berita yang ada pada suatu media di setiap hari atau waktu tertentu menjadi kategori yang telah ditentukan. Hasil pengklasifikasian berita, kemudian diolah dan disajikan dalam bentuk grafik serta dihitung nilai keberpihakannya untuk mengetahui keberpihakan media *online* kepada salah satu pasangan Presiden dan Wakil Presiden. Salah satu metode pengklasifikasian adalah dengan menggunakan metode *Naive Bayes Classifier* (NBC). NBC adalah klasifikasi statistik yang bisa memprediksi probabilitas sebuah kelas, dan kelebihan dari metode ini adalah tingkat akurasi yang tinggi juga waktu komputasi yang lebih cepat.

Berdasarkan pemaparan di atas, penelitian yang akan dilakukan adalah membangun sistem yang secara otomatis dapat mengklasifikasikan berita-berita yang ada pada media *online* menjadi kategori tertentu, serta sistem yang dapat menganalisa dan menentukan keberpihakan media *online* kepada pasangan Presiden dan Wakil Presiden berdasarkan waktu tertentu (*trend* waktu). Data berita yang akan digunakan untuk analisa dan menentukan nilai keberpihakan media *online* serta proses klasifikasi berita adalah data berita yang terdapat di portal-portal media online di Indonesia yang berkaitan dengan pemilihan presiden dan Wakil Presiden.

## 2. Landasan Teori

### 2.1 Text Mining

Pengklasifikasian artikel berita secara otomatis bisa dikategorikan sebagai *text mining*. Proses *text mining* dibagi menjadi 3 tahap utama, yaitu proses awal terhadap teks (*text preprocessing*), transformasi teks ke dalam bentuk antara (*text transformation / feature generation*), dan penemuan pola (*pattern discovery*). Manfaat dari *text mining* adalah untuk mencari pola dalam teks, yaitu proses penganalisisan teks guna menyarikan informasi yang bermanfaat untuk tujuan tertentu. Secara umum proses dalam *text mining* dilakukan dalam tiga tahapan yakni pemilihan data sebagai sumber informasi, kemudian dilakukan *preprocessing* dan

analisis, dan yang terakhir adalah hasil dari proses yang menunjukkan interpretasi dari data yang diuji .

### 2.2. Klasifikasi

Klasifikasi teks merupakan sebuah teknik dalam *text mining* yang bertujuan untuk menempatkan teks pada kategori yang sesuai dengan karakteristiknya dari teks tersebut dengan menggunakan aturan-aturan tertentu. Dengan adanya klasifikasi teks, maka dapat memberikan pandangan secara konseptual mengenai cara pengelompokan data yang memiliki peranan penting terhadap dunia nyata. Tujuan dari pengkategorian teks adalah untuk mengklasifikasikan data ke dalam beberapa kategori-kategori tertentu. Tiap dokumen dapat diklasifikasikan dalam beberapa kategori, atau tidak sama sekali. Dengan menggunakan *machine learning*, pembelajaran dapat dilakukan dengan aturan klasifikasi yang sudah ditentukan dan adanya data latih sebagai acuan pembelajaran sehingga dapat melakukan proses pengklasifikasian secara otomatis nantinya.

### 2.3. Text Preprocessing

Sebelum melakukan analisis teks, *preprocessing* harusnya dilakukan terlebih dahulu untuk mengeliminasi kata-kata yang tidak diperlukan. Hal ini bertujuan agar maksud dari kalimat lebih jelas lagi. Tahapan ini bertujuan untuk mempersiapkan teks menjadi data yang akan mengalami pengolahan pada tahapan berikutnya. Tindakan yang dilakukan meliputi tindakan kompleks dan tindakan sederhana. Contoh tindakan yang bersifat kompleks pada tahap ini adalah *part-of-speech* (*pos*) *tagging*, membangkitkan *parse tree*. Contoh tindakan yang bersifat sederhana adalah proses parsing sederhana terhadap teks, yaitu memecah suatu kalimat menjadi sekumpulan kata. Selain itu pada tahapan ini biasanya juga dilakukan *case folding*, yaitu pengubahan karakter huruf menjadi huruf kecil. Tahapan *text preprocessing* bertujuan untuk mempersiapkan *text* menjadi data yang akan mengalami pengolahan pada tahapan berikutnya. Terdapat tiga langkah dalam tahapan ini adalah *Case Folding*, *Tokenizing*, *Filtering*, *Stemming*. *Case Folding* bertujuan untuk mengubah seluruh huruf dari 'a' sampai dengan 'z' dalam dokumen menjadi huruf kecil. Tidak semua dokumen konsisten dengan penggunaan huruf kapital. Maka dari itu *case folding* mengkonversi keseluruhan teks dalam dokumen menjadi huruf kecil. *Tokenizing*, merupakan tahapan untuk memecah kalimat menjadi kata atau biasa disebut token. Selain itu, tahap ini juga bertujuan untuk membuang beberapa karakter yang dianggap sebagai tanda baca. Dalam proses *tokenizing* juga dilakukan pemisahan dokumen menjadi kata dasar, menghapus awalan, penyisipan, akhiran dan duplikasi. *Filtering*, tahap *filtering* adalah tahap mengambil kata-kata penting

dari hasil *tokenizing* menggunakan algoritma *algoritme stoplist* (membuang kata yang tidak penting) dan *wordlist* (menyimpan kata yang penting). Tujuan umum dari proses ini adalah untuk mendapatkan representasi dasar dari dokumen yang diujikan. *Stemming*, stemming merupakan suatu proses yang mentransformasikan kata-kata yang terdapat dalam suatu dokumen ke kata-kata akarnya (*root word*) dengan menggunakan aturan-aturan tertentu. Sebagai contoh, kata bersama, kebersamaan, menyamai, akan diubah menjadi kata dasar “sama”. Pencarian kata dasar pada setiap bahasa berbeda-beda, untuk bahasa Indonesia sendiri *stemming* menghilangkan imbuhan di awal, imbuhan berada ditengah, imbuhan berada diakhir, ataupun imbuhan di awal dan di akhir dari kata.

2.4 Naïve Bayes Classifier

*Naive Bayes* didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai *output*. Dengan kata lain, diberikan nilai *output*, probabilitas mengamati secara bersama adalah produk dari probabilitas individu.

Keuntungan penggunaan *Naive Bayes* adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (*Training Data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. *Naive Bayes* sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

*Naive Bayes* adalah salah satu algoritma pembelajaran induktif yang paling efektif untuk machine learning dan data mining. Performa *Naive Bayes* yang kompetitif dalam proses klasifikasi walaupun menggunakan asumsi independen atribut tersebut dilanggar performa pengklasifikasian *naive bayes* cukup tinggi, hal ini dibuktikan pada berbagai penelitian empiris .

Secara umum, proses klasifikasi dengan menggunakan metode *Naive Bayes* dapat digambarkan pada persamaan :

$$P(C_j|d) = \frac{P(d|C_j)P(C_j)}{P(d)} \tag{1}$$

Keterangan :

- P(C<sub>j</sub>|d) = probabilitas posterior bersyarat (*conditional probability*) suatu kejadian H terjadi jika diberikan *evidence*
- P(C<sub>j</sub>) = probabilitas awal (prior) kejadian H terjadi tanpa memandang *evidence* apapun.
- P(d) = probabilitas awal (prior) *evidence* X terjadi tanpa memandang kejadian/*evidence* yang lain.
- P(d|C<sub>j</sub>) = probabilitas sebuah *evidence* X terjadi akan mempengaruhi hipotesis H

Dengan pendekatan *Naive Bayes* yang mengasumsikan bahwa tiap-tiap kata didalam setiap kategori adalah tidak bergantung satu sama lain, maka perhitungannya dapat disederhanakan menjadi seperti berikut ini:

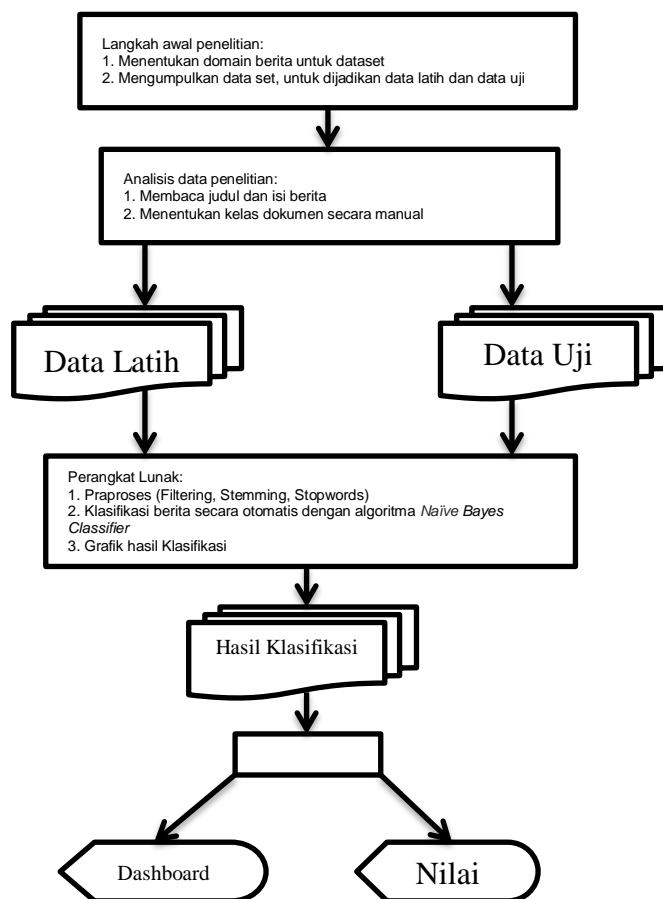
$$P(C_j|d) = \frac{N_{ic} + 1}{N_c + V} \tag{2}$$

Keterangan :

- N<sub>ic</sub> = Jumlah dokumen latih dengan fitur atribut W<sub>i</sub> dengan kategori C<sub>j</sub>
- N<sub>c</sub> = Jumlah dokumen dari kategori C<sub>j</sub>
- V = Jumlah kategori

3. Metode

3.1 Metodologi



Gambar 1. Alur Metodologi Penelitian

Pada Gambar 1 menjelaskan langkah yang akan dilakukan dalam merancang sistem untuk keberpihakan media *online* menggunakan metode *Naive Bayes Classifier* dari tahapan penelitian adalah sebagai berikut, yaitu : Menentukan domain, Domain *berita* yang digunakan dalam penelitian ini adalah pilpres. *berita* yang mengandung kata kunci “pilpres” akan otomatis dikumpulkan. Pengumpulan *Dataset*, Data dalam penelitian ini bersumber dari *website*

tempo.com dan medcom.com Pengambilan data menggunakan proses crawling. *Preprocessing*, Proses yang akan dilakukan pada tahap ini yaitu : *Cleaning*, yaitu proses membersihkan data *berita* dari kata-kata yang tidak mendeskripsikan sesuatu untuk mengurangi *noise*, selain penghapusan kata, dilakukan penghapusan tanda baca seperti koma (,) dan titik (.). *Case folding*, yaitu penyeragaman bentuk huruf, sehingga data hanya menjadi huruf latin dari a sampai z. *Tokenizing*, yaitu proses memecah sebuah kalimat menjadi kata. Klasifikasi, pada tahap ini akan dilakukan pengelompokkan *tweet* ke dalam kelas positif, negatif dan netral. Algoritma yang digunakan yaitu *Naïve Bayes Classifier*. Setelah diklasifikasi maka akan diuji kualitas dari model yang dihasilkan dengan menghitung nilai akurasi. Dan data hasil kklasifikasi kemudian disajikan dalam bentuk grafik

3.2 Studi Literatur

Pada tahapan ini penelitian dilakukan dengan cara mempelajari berbagai literatur melalui pengumpulan dokumen-dokumen, referensi-referensi, buku-buku, sumber dari internet, yang mendukung dan diperlukan untuk merancang sistem yang berkaitan dengan penulisan skripsi yang dilakukan.

3.3 Dataset Berita dan Pengkategorian Berita

Untuk mendapatkan data yang akan digunakan sebagai data uji adalah menggunakan proses *crawling*. Proses *crawling* data yang dibutuhkan dibatasi dengan waktu yang telah ditentukan. Data dengan jangka waktu yang telah ditentukan dari proses *crawling* kemudian akan menuju pada tahap *text processing*, yang bertujuan untuk memudahkan dalam proses pengklasifikasian data. Data berita pada masing-masing media, kemudian dikategorikan menjadi 5 kategori berita. Kategori berita dapat dilihat pada Tabel 1.

Tabel 1. Kategori Berita

No	Kategori	Keterangan
1	Positif A	Berita positif untuk pasangan A
2	Positif B	Berita positif untuk pasangan B
3	Negatif A	Berita negatif untuk pasangan A
4	Negatif B	Berita negatif untuk pasangan B
5	Netral	Berita positif untuk pasangan A dan B

Sebanyak 600 dokumen yang diperoleh dari proses *crawling* di *website* portal berita dibagi menjadi 70% sebagai dokumen latih dan 30% sebagai dokumen uji. Klasifikasi dengan *Naive bayes classifier*. Dalam metode *Naive Bayes classifier* dilakukan proses pengklasifikasian teks berdasarkan data latih yang sebelumnya telah ada didalam database sistem. Penelitian terkait *Naive Bayes Classifiers* melalui beberapa tahap. Evaluasi dilakukan dengan cara membandingkan antara hasil klasifikasi *Naive Bayes Classifier* dan hasil klasifikasi manual oleh petugas spkt polrestabes semarang serta dihitung akurasi dengan menggunakan *confussion matrix*.

4 Hasil dan Pembahasan

4.1 Pembagian Dataset

Dari seluruh *dataset* berita yang sudah dikategorikan, diambil sebagian kemudian dibagi menjadi 4 sub dataset dengan rincian sesuai pada Tabel 2.

Tabel 2. Pembagian Dataset

Sub Dataset	Jumlah Berita				
	Positif A	Negatif A	Netral	Positif B	Negatif B
A	6	6	6	6	6
B	6	6	6	6	6
C	6	6	6	6	6
D	6	6	6	6	6
Total	24	24	24	24	24

Dalam proses pengujian, dataset berita yang telah terbagi ke dalam sub *dataset*, digunakan sebagai data testing sekaligus data *training*. Dalam cara pengujian, data testing sebagai data uji adalah data yang belum memiliki kategori berita, sedangkan data *training* adalah data berita yang sudah memiliki kategori.

Untuk mengkategorikan data testing yang belum memiliki kategori, akan diujikan ke data training dengan menggunakan metode *Naïve Bayes Classifier* dan menentukan nilai *Recall*, *Precision*, dan *Accuracy*. Pengujian data *training* ke data *testing*, dengan rincian sesuai pada Tabel 3.

Tabel 3. Cara Pengujian

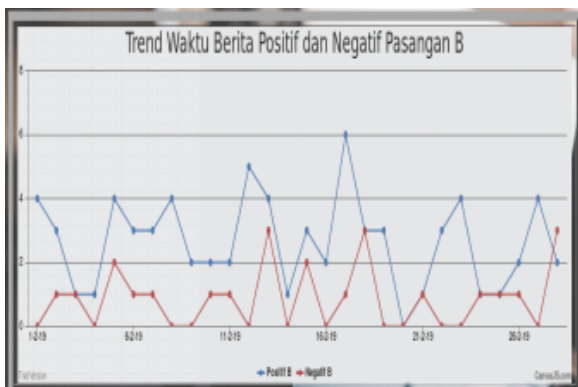
Pengujian Ke -	Data Training (Sub Dataset)	Jumlah Data	Data Testing (Sub Dataset)	Jumlah Data
1	A	30	B, C, D	90
2	B	30	A, C, D	90
3	C	30	A, B, D	90
4	D	30	A, B, C	90

4.2 Grafik Trend Waktu



Gambar 2. Grafik *Trend Waktu* Berita Positif dan Negatif Pada Pasangan A di Media A

Pada Gambar 2 merupakan Grafik *Trend Waktu* jumlah berita positif dan negatif untuk pasangan A di media A.



Gambar 3. Grafik *Trend Waktu* Berita Positif dan Negatif Pada Pasangan B di Media A

Pada Gambar 3 merupakan Grafik *Trend Waktu* jumlah berita positif dan negatif untuk pasangan B di Media A sesuai dengan dengan Tabel 1. Untuk jumlah berita positif dan negatif untuk masing-masing kategori pada masing-masing pasangan yang ada di media A sesuai dengan Tabel 4.

Tabel 4. Jumlah Berita Media A

Jumlah Berita			
Positif A	Negatif A	Positif B	Negatif B
95 Berita	23 Berita	74 Berita	24 Berita

Untuk menghitung nilai keberpihakan pada pasangan A di Media A menggunakan rumus :

$$\begin{aligned} \text{Keberpihakan A} &= \sum \text{Berita Positif A} - \sum \text{Berita Negatif A} \\ &= 95 \text{ Berita} - 23 \text{ Berita} \\ &= 72 \text{ Berita} \end{aligned}$$

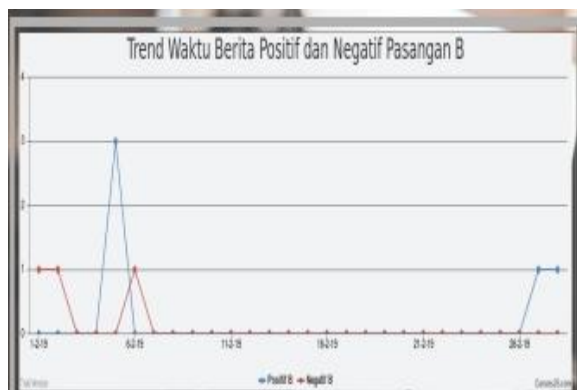
Untuk menghitung nilai keberpihakan pada pasangan B di Media A menggunakan rumus :

$$\begin{aligned} \text{Keberpihakan B} &= \sum \text{Berita Positif B} - \sum \text{Berita Negatif B} \\ &= 74 \text{ Berita} - 24 \text{ Berita} \\ &= 50 \text{ Berita} \end{aligned}$$



Gambar 4. Grafik *Trend Waktu* Berita Positif dan Negatif Pada Pasangan B di Media B

Pada Gambar 4 merupakan Grafik *Trend Waktu* jumlah berita positif dan negatif untuk pasangan A di media B.



Gambar 5. Grafik *Trend Waktu* Berita Positif dan Negatif Pada Pasangan B di Media B

Pada Gambar 5 merupakan Grafik *Trend Waktu* jumlah berita positif dan negatif untuk pasangan B di media B.

Untuk masing-masing kategori pada masing-masing pasangan yang ada di media B sesuai dengan Tabel 5.

Tabel 5. Jumlah Berita Media B

Jumlah Berita			
Positif A	Negatif A	Positif B	Negatif B
65 Berita	1 Berita	12 Berita	16 Berita

Untuk menghitung nilai keberpihakan pada pasangan A di Media A menggunakan rumus :

$$\text{Keberpihakan A} = \sum \text{Berita Positif A} - \sum \text{Berita Negatif A}$$

$$= 65 \text{ Berita} - 1 \text{ Berita}$$

$$= 64 \text{ Berita}$$

Untuk menghitung nilai keberpihakan pada pasangan B di Media A menggunakan rumus :

$$\text{Keberpihakan B} = \sum \text{Berita Positif B} - \sum \text{Berita Negatif B}$$

$$= 12 \text{ Berita} - 16 \text{ Berita}$$

$$= -4 \text{ Berita}$$

### 4.3. Accuracy, Precision, dan Recall

*Accuracy* merupakan tingkat ketepatan antara data atau informasi yang ditemukan oleh sistem dibandingkan dengan jumlah data sebenarnya. Nilai *Accuracy* tertinggi adalah 1, yang berarti seluruh dokumen yang ditemukan adalah relevan.

$$\text{Accuracy} = \frac{\sum \text{benar sistem}}{\sum \text{data sebenarnya}} \quad (3)$$

*Precision* merupakan tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. Nilai *Precision* tertinggi adalah 1, yang berarti seluruh dokumen yang ditemukan adalah relevan.

$$\text{Precision} = \frac{\sum \text{ditemukan sistem}}{\sum \text{data sistem}} \quad (4)$$

*Recall* merupakan tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. Nilai *Recall* tertinggi adalah 1, yang berarti seluruh dokumen dalam koleksi berhasil ditemukan.

$$\text{Recall} = \frac{\sum \text{ditemukan sistem}}{\sum \text{data sebenarnya}} \quad (5)$$

Hasil Pengujian *Accuracy* pada sistem dengan menggunakan metode *Naïve Bayes Classifier* sesuai pada Tabel 6.

Tabel 6. Pengujian *Accuracy*

Pengujian Ke -	Data Traing (Sub Dataset)	Jumlah Data	Data Testing (Sub Dataset)	Jumlah Data	Hasil Accuracy
1	A	30	B, C, D	90	27%
2	B	30	A, C, D	90	50%
3	C	30	A, B, D	90	40%
4	D	30	A, B, C	90	33%

Hasil Pengujian *Precision* pada sistem dengan menggunakan metode *Naïve Bayes Classifier* sesuai pada Tabel 7.

Tabel 7. Pengujian *Precision*

Pengujian Ke -	Hasil Pengujian <i>Precision</i>				
	Positif A	Positif B	Netral	Negatif A	Negatif B
1	17%	50%	50%	17%	0%
2	17%	50%	67%	50%	67%
3	50%	50%	17%	33%	50%
4	33%	33%	33%	50%	17%

Hasil Pengujian *Precision* pada sistem dengan menggunakan metode *Naïve Bayes Classifier* sesuai pada Tabel 8.

Tabel 8. Pengujian *Recall*

Pengujian Ke -	Hasil Pengujian <i>Recall</i>				
	Positif A	Positif B	Netral	Negatif A	Negatif B
1	33%	25%	27%	50%	0%
2	17%	75%	44%	75%	57%
3	60%	100%	33%	18%	38%
4	33%	50%	25%	38%	25%

## 5 Kesimpulan dan Saran

Berdasarkan hasil penelitian dan pengujian tingkat akurasi yang telah dilakukan dapat ditarik kesimpulan bahwa dari hasil pengujian dengan Algoritma *Naïve Bayes Classifier* dapat digunakan untuk mengklasifikasikan kategori berita. Pada analisa keberpihakan media online berdasarkan trend waktu, dapat diambil kesimpulan bahwa Media A maupun media B berpihak kepada pasangan A dengan Nilai Keberpihakan 72 berita positif untuk pasangan A di Media A, dan 64 Berita Positif untuk pasangan A di Media B. Sedangkan untuk pasangan B, nilai keberpihakan di Media A adalah 50 berita positif dan di Media B adalah -4 berita positif atau 4 berita negatif. Nilai perhitungan *accuracy* dari pengujian data testing dan data training dengan pembagian merata untuk setiap kategori di setiap sub dataset adalah pada besaran 27%, 50%, 40%, dan 33%.

Saran yang dapat diberikan dari hasil penelitian untuk pengembangan sistem ini ke depan bahwa penggunaan fitur dan tahap *preprocessing* yang berbeda untuk membandingkan tingkat akurasi. Pengujian metode menggunakan *dataset* yang berbeda untuk membandingkan tingkat akurasi. Optimasi proses pada sistem untuk meningkatkan performa dan memaksimalkan hasil pengujian sistem.

### Daftar Pustaka

Ashadi Siregar. (2006): *Etika Komunikasi, Pustaka Book Publisher.*  
 Asep Syamsul M. Romli. (2012): *Jurnalistik Online: Panduan Mengelola Media Online.* Bandung. Nuansa Cendikia

- Even, Y. & Zohar. (2002): *Introduction to Text Mining*, University of Illionis.
- F Rahutomo, ZHR Adi, IF Rozi, PY Saputra (2018) : *Implementasi Text Mining Pada Website/Blog Di Internet Untuk Menilai Kinerja Suatu Organisasi*, INOVTEK Polbeng-Seri Informatika
- F. Rahutomo, P. Y. Saputra, and M. A. Fidyawan. (2018): *Implementasi Twitter Sentiment Analysis Untuk Review Film Menggunakan Algoritma Support Vector Machine*, JIP, vol. 4, no. 2, p. 93, Feb.g. Thir.
- Garcia, E. (2005): *Document Indexing Tutorial for Information Retrieval Students and Search Engine Marketers*. Available: <http://www.miislita.com/information-retrieval-tutorial/indexing.html>
- Hardiyanto, Erik., Rahutomo, Faisal., dan Puspitasari D., (2016) : *Implementasi K-Nearest Neighbor (KNN) Pada Klasifikasi Artikel Wikipedia*.
- IYR Pratiwi, RA Asmara, F Rahutomo (2017) ; *Study of hoax news detection using naïve bayes classifier in Indonesian language*, 11th International Conference on Information & Communication Technology and System (ICTS).
- Kayser, V. (2016): *Extending the knowlekeysdge base of foresight: The contribution of text mining*. Elsevier, Issue Technological & Sosial Change, pp. 208-215.
- Nokbeh Zaeem, R. M.M.Y.Y. (2016): *Modelling and analysis of identify threat behaviors through text mining of Identify theft stories*. Elsevier, pp.50-63.
- P. Y. Saputra, D. H. Subhi, and F. Z. A. Winatama. (2019): *Implementasi Sentimen Analisis Komentar Channel Video Pelayanan Pemerintah Di Youtube Menggunakan Algoritma Naïve Bayes*, JIP, vol. 5, no. 4, pp. 209-213, Aug.
- Rahayu. (2006): *Potret Pofesionalisme dan Kualitas Pemberitaan Surat Kabar Indonesia. dalam Rahayu (ed.). Menyingkap Profesionalisme Kinerja Surat Kabar di Indonesia*. Pusat Kajian Media dan Budaya Populer, Dewan Pers dan Departemen Komunikasi dan Informasi.
- Sriram, B. (2010): *Short Text clasification in twitter to improve information filtering*. Preceeding of 33<sup>rd</sup> international ACM SIGIR conference on Research and development in information retrieval, Volume SIGIR'10.
- Thirafi, S.F.M., Rahutomo, F. (2018): *Implementation of Naïve Bayes Classifier Algorithm to Categorize Indonesian Song Lyrics Based on Age*, International Conference on Sustainable Information Engineering And Technology (SIET), Malang, Indonesia, pp. 106-109.
- Zhang, & Gao, F. (2011): *An Improvement to NB for Text Classification*. Procedia Engineering, 15, 2160–2164.

