

Perbandingan *Feature Selection Chi-Square* Dan *Query Expansion Ranking (QER)* Pada Analisis Sentimen Terkait Revitalisasi Monas Menggunakan Metode *Naïve Bayes Classifier*

Roni Hadi Wijaya¹, Gita Indah Marthasari², Christian Sri Kusuma Aditya³

^{1,2,3}Universitas Muhammadiyah Malang

ronihady@gmail.com¹, gita@umm.ac.id², christianskaditya@umm.ac.id³

Abstrak

Monumen Nasional menjadi salah satu destinasi yang cukup populer. Di awal tahun 2020, taman Monumen Nasional dilakukan revitalisasi. Dengan adanya kebijakan tersebut, mengundang banyak kalangan untuk memberikan tanggapan secara langsung maupun lewat social media. Dengan demikian, penelitian ini mencoba menganalisis ulasan dari masyarakat dengan metode *naive bayes classifier* dan membandingkan penggunaan seleksi fitur *chi-square* dan *query expansion ranking* untuk mengoptimalkan kinerja klasifikasi pada penelitian. Hasil klasifikasi *naive bayes classifier* yang didapatkan dengan seleksi fitur *query expansion ranking* sebesar 80% dengan variasi rasio seleksi fitur 75%. Hasil tersebut lebih baik dibandingkan dengan hasil klasifikasi *naive bayes classifier* dengan seleksi fitur *chi-square* menggunakan variasi rasio 75% dengan akurasi 79%.

Kata Kunci: Analisis Sentimen, *Naïve Bayes*, Seleksi Fitur, Python

Abstract

The National Monument is one of the most popular destinations. In early 2020, the revitalization of the National Monument park was carried out. With this policy, many peoples are invited to respond directly or via social media. Thus, this study tries to analyze reviews from the public with the *Bayes-naïve classification method* and compares the use of *feature selection chi-square* and the *query expansion ranking* to improve classification performance in the study. The results of the classification of the *naive bayes classifier* are obtained by *feature selection the query expansion ranking* of 80% with a variation of the *feature selection ratio* of 75%. These results are better than the results of the *naive bayes classifier* with *feature selection the chi-square* using a *ratio variation* of 75% with an accuracy of 79%.

Keywords: *Sentiment analysis, Naïve Bayes, Feature Selection, Python*

1. Pendahuluan

Jakarta (DKI Jakarta) merupakan Ibu Kota Negara Republik Indonesia dan Ibu Kota Provinsi yang terletak di pesisir bagian barat laut di Pulau Jawa. Menurut BPS (Badan Pusat Statistik) Jakarta mempunyai luas 662,33 km²[1] dengan jumlah penduduk sebesar 10.57 juta orang[2]. Tugu Monas (Tugu Monumen Nasional) merupakan suatu monumen bersejarah yang terletak di Kota Jakarta. Monumen ini dibangun untuk mengenang perjuangan rakyat Indonesia melawan penjajahan Belanda. Tugu Monas mempunyai tinggi 132m ditambah dengan adanya mahkota yang berbentuk lidah api dengan lapisan lembaran emas membuat Monas terlihat megah.

Tugu Monas menjadi salah satu destinasi pariwisata yang cukup populer di Kota Jakarta. Di awal tahun 2020, Taman Monumen Nasional (Monas) kembali dilakukan revitalisasi, dimana sebelumnya kawasan Monas seringkali dilakukan revitalisasi. Namun, kebijakan revitalisasi yang digaungkan oleh Gubernur DKI Jakarta Anies Baswedan kali ini sempat mendapat penolakan dari berbagai pihak [3]. Begitu pula juga dikabarkan akan diselenggarakan balap Formula E di kawasan Monumen Nasional, dimana hal ini dikhawatirkan akan merusak kawasan Monas yang merupakan cagar budaya. Meski demikian, proyek revitalisasi ini tetap berjalan [4].

Dengan adanya kejadian tersebut, mengundang banyak kalangan untuk memberikan tanggapan mengenai revitalisasi di kawasan Monas yang di harapkan tidak merusak kawasan

Monas. Tanggapan tersebut diungkapkan secara langsung maupun lewat media sosial, salah satunya Twitter. Media sosial menjadi salah satu wadah bagi semua orang untuk berkomunikasi, menyampaikan pendapat atau sekedar untuk mengekspresikan diri. Menurut laman web sindo pengguna media sosial di Indonesia mencapai 150 juta pengguna. Perkembangan media sosial Twitter di Indonesia sangat pesat, Twitter menjadi salah satu media sosial yang mempunyai pengguna sebanyak 6.43 juta pengguna[5]. Setiap pengguna dapat membagikan Tweet dengan maksimal jumlah karakter sebanyak 280 karakter yang dapat dikombinasikan dengan foto, video ataupun tautan.

Algoritma *Naïve Bayes Classifier* merupakan algoritma pengklasifikasian statistik yang dapat memprediksi nilai probabilitas anggota dari suatu class. Namun, ketidakakuratan data yang dimiliki dapat menyebabkan adanya kesalahan dalam proses pengklasifikasian datanya. Walaupun sebenarnya algoritma *Naïve Bayes* merupakan salah satu algoritma yang dapat meminimalkan kesalahan dari pada algoritma pengklasifikasian yang lain. Pada penelitian sebelumnya, mengenai sentimen analisis mengenai kasus anti-LGBT di Indonesia dengan menggunakan metode *Naïve Bayes Classifie*, *Decision Tree* dan *Random Forest Algorithm*. Hasil dari penelitian tersebut diketahui dengan metode *Naïve Bayes* sebesar 86,43% sedangkan *Decision Tree* dan *Random Forest Algorithm* sebesar 82,91% [6].

Dalam penelitian mengenai implementasi *Naïve Bayes* terhadap analisis sentiment opini film pada twitter yang dilakukan oleh Fajar Ratnawati, dkk. Penelitian tersebut dilakukan untuk mengklasifikasikan kecenderungan opini pengguna twitter mengenai film apakah cenderung terhadap opini positif atau negatif. Diketahui hasil yang didapatkan pada penelitian tersebut mendapat akurasi 90% [7].

Feature Selection sangat efektif digunakan karena dapat menghilangkan data yang tidak relevan sehingga dapat mengoptimalkan kinerja metode klasifikasi yang digunakan menjadi lebih baik, efektif dan efisien. Metode *feature selection* diantaranya *Chi Square*, *Query Expnasion Ranking (QER)*, *PSO*, *Information Gain* dsb. Pada penelitian sebelumnya mengenai analisis sentiment *tweet* berbahasa sunda dengan menggunakan metode pengklasifikasian *Naïve Bayes* dan *feature selection Chi Square*. *Chi Square* digunakan untuk mengurangi dimensi data yang tinggi agar data lebih relevan dalam melakukan klasifikasi dengan *Naïve Bayes*. Hasil dari penelitian ini dengan menerapkan *feature selection Chi Square* akurasi sebesar 78,48% [8]. Penelitian lain yang dilakukan oleh Mokhammad Rifqi Tsani, dkk tentang analisis sentiment review transportasi dengan menggunakan algoritma *Support vector Machine* dan *feature selection Chi Square*. Hasil yang di dapatkan akurasi sebesar 88,29% [9].

Penelitian sebelumnya, mengenai sentiment analisis untuk mengetahui opini tentang kurikulum 2013 di media sosial. Penelitian ini menggunakan metode klasifikasi *K-Nearest Neighbor* dan *feature selection Query Expansion Ranking (QER)*. *Query Exoansion Ranking* digunakan setelah dilakukan pembobotan *TF-IDF*, hal ini dimaksudkan agar data lebih efektif. Hasil dari penelitian tersebut diketahui sebesar 96,36% dengan menggunakan rasio *QER* 50%. Pengujian ini membuktikan bahwa *feature selection QER* dapat meningkatkan akurasi [10].

Penelitian lain tentang sentiment analisis pariwisata di kota Malang, yang bertujuan untuk mengetahui opini mengenai pariwisata yang ada di kota Malang. Penelitian ini menggunakan metode *Naïve Bayes Classifier* dengan *feature selection Query Expansion Ramking (QER)*. Hasil dari penelitian ini diketahui akurasi terbaik sebesar 86,6% [11].

Dalam penelitian kali ini akan melakukan klasifikasi pada opini masyarakat mengenai revitalisasi kawasan Monumen Nasional (Monas) pada awal tahun 2020 dengan menggunakan metode *Naïve Bayes Classifier*. *Feature selection* yang akan digunakan untuk mengurangi data yang kurang relevan dan mengoptimalkan kinerja klasifikasi pada penelitian ini yaitu *Chi Square* dan *Query Expansion Ranking (QER)*. Penggunaan kedua *feature selection Chi Square* dan *Query Expansion Ranking (QER)* yang membedakan penelitian ini dari penelitian sebelumnya..

2. Metode Penelitian

2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data *tweet* mengenai topik kebijakan Revitalisasi Monas di Kota Jakarta pada awal tahun 2020. Proses *crawling* data dilakukan dengan menggunakan *library Twtterscraper* pada *Python* dengan kata kunci “revitalisasi monas” pada tanggal 01 Januari 2020 – 31 Juli 2020 dengan jumlah data sebanyak 1000 data yang terdiri dari 196 data berlabel positif dan 804 data berlabel negatif. Labeling dilakukan oleh 3 orang annotator.

Untuk menggunakan *library Twitterscraper*, diharuskan untuk menginstall terlebih dahulu dengan memasukkan *keyword* “*pip install twitterscraper*” pada direktori *file Python* yang digunakan.

2.2 Preprocessing

Tahap preprocessing merupakan tahap awal setelah mengumpulkan data yang dibutuhkan dengan teknik *crawling* data. Pada preprocessing, data diproses untuk menghilangkan kata yang *noise*. Berikut beberapa tahapan yang dilakukan pada proses *preprocessing* :

2.2.1 Case folding

Pada tahap case folding, setiap kalimat yang mengandung kata – kata huruf besar (*uppercase*) akan diubah menjadi huruf kecil (*lower case*)[7]. Hasil dari tahap case folding adalah semua huruf kecil yang seragam dari ‘a’ sampai dengan ‘z’.

2.2.2 Filtering

Pada tahap ini pastinya *tweet* mengandung banyak kata – kata yang diperlukan dan tidak diperlukan dalam proses klasifikasi nanti, maka setiap kata – kata yang tidak penting akan dihapus sehingga hasilnya hanya kata – kata yang diperlukan saja. Terdapat dua proses dalam tahap *filtering*, yaitu sebagai berikut :

1. Punctuation Removal

Punctuation Removal adalah proses yang bertujuan untuk menghapus tanda baca dan angka yang ada pada *tweet*[12].

2. Stopword Removal

Stopwords Removal merupakan proses menghilangkan kata – kata yang tidak berpengaruh pada deskripsi. Kata – kata yang terdapat pada *tweet* akan dihapus sesuai dengan daftar pada *stopword* misalnya kata penghubung “atau”, “dan” atau kata – kata yang tidak berpengaruh pada tahapan klasifikasi [13].

2.2.3 Tokenizing

Pada tahap tokenizing, data yang sebelumnya berupa kalimat akan dipecah menjadi kata per-kata, agar selanjutnya dapat dilakukan pembobotan pada setiap katanya[13].

2.2.4 Stemming

Stemming merupakan proses untuk mendapatkan kata dasar dari suatu kata dalam kalimat dengan cara memisahkan kata dari kata dasar dan imbuhan nya baik awalan dan akhiran. Seperti, diperlukan memerlukan, akan di *stem* ke kata dasarnya “perlu”[14].

2.3 Term Frequency Inverse Document Frequency

TF-IDF (*Term Frequency – Invers Document Frequency*) merupakan proses memberikan nilai pada *term* dengan melakukan perhitungan nilai *term frequency (TF)* lalu melakukan perhitungan *invers document frequency (IDF)*. Pada tahap TF digunakan untuk mencari nilai dari kemunculan kata dalam suatu dokumen. Lalu tahap IDF digunakan untuk mencari nilai kemunculan dari kata pada keseluruhan dokumen, nilai IDF berbanding terbalik dengan TF, semakin banyak kata yang muncul maka nilai IDF akan semakin kecil. Untuk menghitung TF – IDF dari kata, menggunakan Persamaan 1.

$$TF - IDF(t, d, D) = tf(t, d) \times idf(t, D) \quad (1)$$

2.4 Feature Selection

Feature selection yang akan digunakan untuk mengurangi data yang kurang relevan dan mengoptimalkan kinerja klasifikasi pada penelitian. Terdapat dua seleksi fitur, yaitu seleksi fitur *Chi Square* dan seleksi fitur *Query Expansion Ranking*.

2.4.1 Chi Square

Chi Square merupakan suatu konsep statistika yang digunakan untuk menguji nilai kesamaan antara *term* dan kategorinya. *Chi Square* digunakan untuk membersihkan fitur – fitur yang tidak sesuai agar proses klasifikasi dapat lebih optimal[8]. Persamaan dari *Feature Selection Chi Square* pada Persamaan 2.

$$x^2(t, c) = \frac{N(A \times D - C \times B)^2}{(A + C)(B + D)(A + B)(C + D)} \quad (2)$$

2.4.2 Query Expansion Ranking (QER)

Query Expansion Ranking (QER) merupakan salah satu metode seleksi fitur yang digunakan untuk mengurangi kompleksitas komputasi dalam analisis sentimen dan memberikan skor pada setiap kata [11].s Persamaan 3 *Query Expansion Ranking (QER)* untuk menghitung peluang fitur f pada dokumen kelas positif.

$$p_f = \frac{df_+^f + 0,5}{n^+ + 1,0} \quad (3)$$

Persamaan 4 untuk menghitung peluang fitur f pada dokumen kelas negatif.

$$q_f = \frac{df_-^f + 0,5}{n^- + 0,5} \quad (4)$$

Setelah mendapatkan jumlah dokumen positif dan dokumen negatif, selanjutnya akan menghitung nilai dari *Query Expansion Ranking* pada Persamaan 5.

$$Score_f = \frac{|p_f + q_f|}{|p_f - q_f|} \quad (5)$$

2.5 Klasifikasi Naïve Bayes

Menurut Wu dan Kumar bahwa *Naïve Bayes Classifier* masuk dalam 10 algoritma terbaik dalam data mining sebagai sebuah metode klasifikasi yang populer [6]. *Naïve Bayes Classifier* adalah pengklasifikasian dengan menggunakan perhitungan probabilitas. Dalam metode ini dilakukan dua proses yaitu proses pelatihan data dengan membuat model klasifikasi, yang kedua proses pengujian dengan menggunakan data uji yang dimasukkan kedalam model klasifikasi yang telah dibuat menggunakan proses pelatihan data.

Pada proses klasifikasi algoritma akan mencari probabilitas tertinggi dari semua kategori dokumen yang diujikan. Berikut Persamaan 6 yang digunakan dalam proses klasifikasi.

$$P(c|d) = P(c) \prod_{i=1}^n P(w_i|c) \quad (6)$$

2.6 Pengujian

Pengujian dilakukan dengan membagi data latih dan data uji pada tahap klasifikasi di setiap skenario pengujian. Setelah itu akan dilakukan penghapusan fitur pada setiap seleksi fitur dengan menggunakan pembagian data latih dan data uji yang mempunyai akurasi terbaik. Persamaan 7, Persamaan 8, Persamaan 9, dan Persamaan 10 untuk pengukuran akurasi diukur dengan *confusion matrix*.

Tabel 1. *Confussion Matrix*

	Positif	Negatif
Positif	True Positif	False Positif
Negatif	False Negatif	True Negatif

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$Presisi = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F - measure = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (10)$$

3. Hasil Penelitian dan Pembahasan

Pembahasan dan penelitian terdapat dua proses klasifikasi. Pertama klasifikasi dengan metode *naïve bayes classifier* berbasis *chi square* dan yang kedua menggunakan metode *naïve bayes classifier* berbasis *query expansion ranking*. Masing – masing proses klasifikasi metode *naïve bayes classifier* berbasis *chi square* dan klasifikasi metode *naïve bayes classifier* berbasis *query expansion ranking* melakukan 2 tahap pengujian dengan masing – masing tahap pengujian terdiri dari 3 skenario pengujian.

Tahap pertama, yaitu pengujian dengan membagi data latih dan data uji yang terdiri dari skenario pertama yaitu dengan pembagian data uji sebesar 25% dan data latih sebesar 75%. Skenario kedua dengan pembagian data uji 50% dan data latih 50%. Skenario ketiga dengan pembagian data latih 75% dan data uji sebesar 25%. Setelah melakukan pengujian tahap 1, akurasi terbaik dari pengujian tersebut akan digunakan lagi untuk pengujian tahap 2 dengan mengurangi fitur seleksi. Pada pengujian tahap 2 ini akan dilakukan dengan 3 skenario, yaitu skenario pertama dengan mengurangi fitur sebanyak 25% dari fitur atau *term* keseluruhan, skenario kedua dengan 50% fitur atau *term* yang di kurangi, dan skenario ketiga 75% fitur atau *term* yang di kurangi

3.1 Klasifikasi Naïve Bayes Classifier Menggunakan Chi – Square

Pada Gambar 1, skenario pertama dengan data latih sebesar 25% atau 250 data dan data uji sebesar 75% atau 750 data. Dihasilkan nilai pada kolom *True Positive* sebesar 158, *False Positive* sebesar 0, *False Negative* sebesar 592, dan *True Negative* sebesar 0.

	precision	recall	f1-score	support
Negatif	0.79	1.00	0.88	592
Positif	0.00	0.00	0.00	158
accuracy			0.79	750
macro avg	0.39	0.50	0.44	750
weighted avg	0.62	0.79	0.70	750
[[592 0]				
[158 0]]				

Gambar 1. Confussion Matrix

Pada Gambar 2, skenario kedua dengan data latih sebesar 50% atau 500 data dan data uji sebesar 50% atau 500 data. Dihasilkan nilai pada kolom *True Positive* sebesar 106, *False Positive* sebesar 1, *False Negative* sebesar 393, dan *True Negative* sebesar 0.

	precision	recall	f1-score	support
Negatif	0.79	1.00	0.88	393
Positif	1.00	0.01	0.02	107
accuracy			0.79	500
macro avg	0.89	0.50	0.45	500
weighted avg	0.83	0.79	0.70	500
[[393 0]				
[106 1]]				

Gambar 2. Confussion Matrix

Pada Gambar 3, skenario ketiga dengan data latih sebesar 75% atau 750 data dan data uji sebesar 25% atau 250 data. Dihasilkan nilai pada kolom *True Positive* sebesar 54, *False Positive* sebesar 1, *False Negative* sebesar 196, dan *True Negative* sebesar 0.

	precision	recall	f1-score	support
Negatif	0.78	1.00	0.88	196
Positif	0.00	0.00	0.00	54
accuracy			0.78	250
macro avg	0.39	0.50	0.44	250
weighted avg	0.61	0.78	0.69	250
[[196 0]				
[54 0]]				

Gambar 3. Confussion Matrix

Berdasarkan hasil yang didapatkan dari *confusion matrix* skenario pertama sampai ketiga. Maka, didapatkan hasil dari evaluasi klasifikasi dari nilai *accuracy*, *precision*, *recall* dan *f-measure*. Berikut nilai dari hasil evaluasi klasifikasi tersebut, pada Tabel 2.

Tabel 2. Evaluasi Klasifikasi

Skenario	Accuracy	Precision	Recall	F-Measure
1	79%	64%	79%	70%
2	79%	83%	79%	70%
3	78%	61%	78%	69%

Dari hasil klasifikasi *naïve bayes classifier* dengan menggunakan *feature selection chi-square* diketahui pada Tabel 2 yang memiliki nilai terbesar dari hasil *accuracy*, *precision*, *recall* dan *f-measure* adalah pada keadaan data 50% data latih dan 50% data uji.

Pada Gambar 4, skenario empat dengan penghapusan fitur sebesar 25% atau 596 fitur dengan keadaan data latih sebesar 50% atau 500 data dan data uji sebesar 50% atau 500 data. Dihasilkan nilai pada kolom *True Positive* sebesar 105, *False Positive* sebesar 0, *False Negative* sebesar 393, dan *True Negative* sebesar 2.

	precision	recall	f1-score	support
Negatif	0.79	1.00	0.88	393
Positif	1.00	0.02	0.04	107
accuracy			0.79	500
macro avg	0.89	0.51	0.46	500
weighted avg	0.83	0.79	0.70	500
[[393 0]				
[105 2]]				

Gambar 4. Confussion Matrix

Pada Gambar 5, skenario 5 dengan penghapusan fitur sebesar 50% atau 1192 fitur dengan keadaan data latih sebesar 50% atau 500 data dan data uji sebesar 50% atau 500 data. Dihasilkan nilai pada kolom *True Positive* sebesar 105, *False Positive* sebesar 2, *False Negative* sebesar 393, dan *True Negative* sebesar 0.

	precision	recall	f1-score	support
Negatif	0.79	1.00	0.88	393
Positif	1.00	0.02	0.04	107
accuracy			0.79	500
macro avg	0.89	0.51	0.46	500
weighted avg	0.83	0.79	0.70	500
[[393 0]				
[105 2]]				

Gambar 5. Confussion Matrix

Pada Gambar 6, skenario 6 dengan penghapusan fitur sebesar 75% atau 1788 fitur dengan keadaan data latih sebesar 50% atau 500 data dan data uji sebesar 50% atau 500 data. Dihasilkan nilai pada kolom *True Positive* sebesar 103, *False Positive* sebesar 4, *False Negative* sebesar 393, dan *True Negative* sebesar 0.

	precision	recall	f1-score	support
Negatif	0.79	1.00	0.88	393
Positif	1.00	0.04	0.07	107
accuracy			0.79	500
macro avg	0.90	0.52	0.48	500
weighted avg	0.84	0.79	0.71	500
[[393 0]				
[103 4]]				

Gambar 6. Confussion Matrix

Berdasarkan hasil yang didapatkan dari *confusion matrix* skenario 4 sampai 6. Maka, didapatkan hasil dari evaluasi klasifikasi dari nilai *accuracy*, *precision*, *recall* dan *f-measure*. Tabel 3 berikut nilai dari hasil evaluasi klasifikasi tersebut.

Tabel 3. Evaluasi Klasifikasi

Skenario	Accuracy	Precision	Recall	F-Measure
4	79%	83%	79%	70%
5	79%	83%	79%	81%
6	79%	84%	79%	81%

3.2 Klasifikasi Naïve Bayes Classifier Menggunakan Query Expansion Ranking

Pada Gambar 7, skenario 1 dengan data latih sebesar 25% atau 250 data dan data uji sebesar 75% atau 750 data. Dihasilkan nilai pada kolom *True Positive* sebesar 647, *False Positive* sebesar 0, *False Negative* sebesar 2035, dan *True Negative* sebesar 0.

	precision	recall	f1-score	support
Negatif	0.76	1.00	0.86	2035
Positif	0.00	0.00	0.00	647
accuracy			0.76	2682
macro avg	0.38	0.50	0.43	2682
weighted avg	0.58	0.76	0.65	2682
Out[5]: array([[2035, 0],				
[647, 0]])				

Gambar 7. Confussion Matrix

Pada Gambar 8, skenario 2 dengan data latih sebesar 50% atau 500 data dan data uji sebesar 50% atau 500 data. Dihasilkan nilai pada kolom *True Positive* sebesar 409, *False Positive* sebesar 0, *False Negative* sebesar 1379, dan *True Negative* sebesar 0.

	precision	recall	f1-score	support
Negatif	0.77	1.00	0.87	1379
Positif	0.00	0.00	0.00	409
accuracy			0.77	1788
macro avg	0.39	0.50	0.44	1788
weighted avg	0.59	0.77	0.67	1788
Out[5]: array([[1379, 0],				
[409, 0]])				

Gambar 8. Confussion Matrix

Pada Gambar 9, skenario 3 dengan data latih sebesar 75% atau 750 data dan data uji sebesar 25% atau 250 data. Dihasilkan nilai pada kolom *True Positive* sebesar 211, *False Positive* sebesar 0, *False Negative* sebesar 683, dan *True Negative* sebesar 0.

```

precision    recall  f1-score   support

Negatif      0.76    1.00    0.87    683
Positif      0.00    0.00    0.00    211

accuracy          0.76    894
macro avg         0.38    0.50    0.43    894
weighted avg     0.58    0.76    0.66    894

Out[5]: array([[683,  0],
               [211,  0]])

```

Gambar 9. Confussion Matrix

Berdasarkan hasil yang didapatkan dari *confusion matrix* skenario pertama sampai ketiga pada tabel. Maka, didapatkan hasil dari evaluasi klasifikasi dari nilai *accuracy*, *precision*, *recall* dan *f-measure*. Berikut Tabel 4 nilai dari hasil evaluasi klasifikasi tersebut.

Tabel 4. Evaluasi Klasifikasi

Skenario	Accuracy	Precision	Recall	F-Measure
1	76%	58%	76%	65%
2	77%	59%	77%	67%
3	76%	58%	76%	66%

Dari hasil klasifikasi *naïve bayes classifier* dengan menggunakan *feature selection query expansion ranking* diketahui yang memiliki nilai terbesar dari hasil *accuracy*, *precision*, *recall* dan *f-measure* adalah pada keadaan 50% data latih dan 50% data uji.

Pada Gambar 10, skenario 4 dengan penghapusan fitur sebesar 25% atau 894 fitur dengan keadaan data latih sebesar 50% atau 500 data dan data uji sebesar 50% atau 500 data. Dihasilkan nilai pada kolom *True Positive* sebesar 393, *False Positive* sebesar 0, *False Negative* sebesar 949, dan *True Negative* sebesar 0.

```

precision    recall  f1-score   support

Negatif      0.71    1.00    0.83    949
Positif      0.00    0.00    0.00    393

accuracy          0.71   1342
macro avg         0.35    0.50    0.41   1342
weighted avg     0.50    0.71    0.59   1342

Out[5]: array([[949,  0],
               [393,  0]])

```

Gambar 10. Confussion Matrix

Pada Gambar 11, skenario 5 dengan penghapusan fitur sebesar 50% atau 1789 fitur dengan keadaan data sebesar 50% atau 500 data dan data uji sebesar 50% atau 500 data. Dihasilkan nilai pada kolom *True Positive* sebesar 387, *False Positive* sebesar 0, *False Negative* sebesar 508, dan *True Negative* sebesar 0.

```

precision    recall  f1-score   support

Negatif      0.57    1.00    0.72    508
Positif      0.00    0.00    0.00    387

accuracy          0.57    895
macro avg         0.28    0.50    0.36    895
weighted avg     0.32    0.57    0.41    895

Out[5]: array([[508,  0],
               [387,  0]])

```

Gambar 11. Confussion Matrix

Pada Gambar 12, skenario 6 dengan penghapusan fitur sebesar 75% atau 2683 fitur dengan keadaan data sebesar 50% atau 500 data dan data uji sebesar 50% atau 500 data. Dihasilkan nilai pada kolom *True Positive* sebesar 0, *False Positive* sebesar 358, *False Negative* sebesar 0, dan *True Negative* sebesar 90.

	precision	recall	f1-score	support
Negatif	0.00	0.00	0.00	90
Positif	0.80	1.00	0.89	358
accuracy			0.80	448
macro avg	0.40	0.50	0.44	448
weighted avg	0.64	0.80	0.71	448

Out[5]: array([[0, 90],
[0, 358]],

Gambar 12. Confussion Matrix

Berdasarkan hasil yang didapatkan dari *confusion matrix* skenario 4 sampai 6. Maka, didapatkan hasil dari evaluasi klasifikasi dari nilai *accuracy*, *precision*, *recall* dan *f-measure*. Berikut nilai dari hasil evaluasi klasifikasi pada Tabel 5.

Tabel 5. Evaluasi Klasifikasi

Skenario	Accuracy	Precision	Recall	F-Measure
4	71%	50%	71%	59%
5	58%	32%	57%	41%
6	80%	64%	80%	71%

Dari hasil perhitungan diatas diketahui jika hasil dari akurasi, presisi, recall, dan f-measure pada penelitian ini menggunakan metode *naïve bayes classifier* berbasis *query expansion ranking* mendapatkan hasil yang paling baik dibandingkan dengan hasil dari klasifikasi *naïve bayes classifier* yang dikombinasikan dengan *chi-square*. Hasil tersebut menggunakan pembagian data 50% data latih dan 50% data uji dengan seleksi fitur sebesar 75%, diketahui hasil *accuracy* sebesar 83% dengan nilai *precision* 70%, nilai *recall* sebesar 86% dan nilai *f-measure* sebesar 76%.

Namun, di setiap skenario pengujian mulai dari skenario 1–6 *chi-square* dan *naïve bayes* dan skenario 1–5 *query expansion ranking* dan *naïve bayes* diketahui nilai *precision*, *recall*, dan *f-measure* pada kelas positif mendapatkan nilai rendah bahkan nol. Hal ini dikarenakan adanya ketidakseimbangan data pada setiap kelas (*imbalance*) yang menyebabkan pada saat klasifikasi, system akan cenderung mengklasifikasi data yang berasal dari kelas positif akan terklasifikasi dengan data kelas negatif sehingga hasil klasifikasi kurang baik.

Adanya data set yang kurang tepat, dimana banyaknya fitur atau data yang digunakan pada data uji tidak termasuk pada kelasnya. Contohnya, ada kata “setuju” pada kelas positif dan juga ada kata “setuju” pada kelas negatif yang sebenarnya berasal dari kata “tidak setuju”. Kata yang sama pada kelas positif dan kelas negatif inilah yang menyebabkan adanya kesalahan dalam proses klasifikasi pada sistem yang berpengaruh pada kinerja sistem, sehingga mempengaruhi tingkat akurasi yang di dapatkan tidak maksimal

Sedangkan pada skenario 6 *query expansion ranking* dan *naïve bayes* nilai *precision*, *recall* dan *f-measure* setelah mengalami peningkatan setelah penghapusan fitur dengan rasio 75% pada kelas positif. Hal ini terjadi karena kelas negatif mengalami penurunan, sehingga data kelas positif dan data kelas negatif tidak seimbang (*imbalance*). Hal ini menyebabkan terjadinya tingkat kesalahan yang tinggi pada sistem dalam menemukan kembali informasi pada opini kelas negatif saat proses klasifikasi atau bahkan tidak dapat menemukan informasi tersebut sehingga nilai pada *precision*, *recall* dan *f-measure* pada kelas negatif mendapatkan nilai nol.

4. Kesimpulan

Berdasarkan pada penelitian yang dilakukan, klasifikasi *naïve bayes classifier* dengan menggunakan *feature selection query expansion ranking* mendapatkan hasil klasifikasi yang paling baik di bandingkan dengan penggunaan *feature selection chi-square* dalam penerapannya.

Penggunaan *feature selection chi – square* dan *query expansion ranking* pada klasifikasi analisis sentimen dengan metode *naïve bayes classifier* diketahui dapat meningkatkan hasil akurasi dari klasifikasi dengan baik, Namun, pada penggunaan *chi – square* sebagai seleksi fitur diketahui dapat meningkatkan hasil akurasi dari klasifikasi, namun tidak terlalu signifikan.

Daftar Notasi

Keterangan Confusion Matrix

- TP : True Positif
 FP : False Positif
 TN : True Negatif
 FN : False Negatif

Referensi

- [1] R. Haryanti, "Luas Ibu Kota Baru di Kalimantan Timur Hampir 3 Kali DKI Jakarta," *Kompas.com*, Jakarta, 26-Aug-2019.
- [2] D. H. Jayani, "Proyeksi Jumlah Penduduk DKI Jakarta 2020," *databoks.katada.co.id*, 2020. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2019/12/07/jumlah-penduduk-dki-jakarta-2020>.
- [3] W. Marison, "Revitalisasi Monas: Dari Ditolak Istana Negara, Mangkrak, hingga Kembali Berjalan," *Kompas.com*, Jakarta, 11-Feb-2020.
- [4] H. Widowati, "Pemerintah Provinsi DKI Jakarta berjanji pembangunan lintasan dan fasilitas pendukung Formula E tidak akan merusak kawasan cagar budaya Monas.," *Katadata.co.id*, Jakarta, 21-Feb-2020.
- [5] "Indonesia Digital 2019 : Media Sosial," *websindo.com*, 07-Mar-2019.
- [6] F. N. Hasan and M. Wahyudi, "Analisis Sentimen Artikel Berita Tokoh Sepak Bola Dunia Menggunakan Algoritma Support Vector Machine Dan Naive Bayes Berbasis Particle Swarm Optimization," *Director*, vol. 15, no. 2, pp. 2017–2019, 2018.
- [7] V. A. Fitri, R. Andreswari, M. A. Hasibuan, V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment Analysis of Social Media Twitter with Case of Anti- LGBT Campaign in Indonesia using Naïve Bayes , Decision Tree , and Random Forest Algorithm," *Procedia Comput. Sci.*, vol. 161, pp. 765–772, 2019.
- [8] F. Ratnawati, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018.
- [9] Y. Cahyono and S. Saprudin, "Analisis Sentiment Tweets Berbahasa Sunda Menggunakan Naive Bayes Classifier dengan Seleksi Feature Chi Squared Statistic," *J. Inform. Univ. Pamulang*, vol. 4, no. 3, p. 87, 2019.
- [10] M. R. Tsani, A. Prima, G. Rupaka, L. Asmoro, P. Keselamatan, and T. Jalan, "Analisis Sentimen Review Transportasi Menggunakan Algoritma Support Vector Machine Berbasis Chi Square," vol. 9, no. 1, pp. 35–39, 2020.
- [11] N. D. Mentari, M. A. Fauzi, and L. Muflikhah, "Analisis Sentimen Kurikulum 2013 Pada Sosial Media Twitter Menggunakan Metode K-Nearest Neighbor dan Feature Selection Query Expansion Ranking," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 8, pp. 2739–2743, 2018.
- [12] S. Mujilawati, "Pre-Processing Text Mining Pada Data Twitter," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2016, no. Sentika, pp. 2089–9815, 2016.
- [13] R. Kurniawan and A. Apriliani, "Analisis Sentimen Masyarakat Terhadap Virus Corona Berdasarkan Opini Dari Twitter Berbasis Web Scraper," *Jurnal INSTEK (Informatika Sains dan Teknologi)*, vol. 5, no. 1. p. 67, 2020.
- [14] D. Wahyudi, T. Susyanto, and D. Nugroho, "Implementasi Dan Analisis Algoritma Stemming Nazief & Adriani Dan Porter Pada Dokumen Berbahasa Indonesia," *J. Ilm. SINUS*, vol. 15, no. 2, pp. 49–56, 2017.
- [15] S. Fanissa, M. A. Fauzi, and S. Adinugroho, "Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking | Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 8, pp. 2766–2770, 2018.