

ANALISIS SENTIMEN OPINI PUBLIK TERHADAP PENGAMBIL ALIHAN TMII OLEH PEMERINTAH DENGAN ALGORITMA NAÏVE BAYES

Ika Amelia¹, Adinda Mutiara², Imam Santoso³

¹STIKOM Cipta Karya Informatika, ²STIKOM Cipta Karya Informatika, ³UTM Jakarta
¹amelia@stikomcki.ac.id, ²dindazani@stikomcki.ac.id, ³imam.santoso@utmj.ac.id

ABSTRAK

Pariwisata memainkan peran penting dalam pembangunan ekonomi suatu negara. Taman Mini Indonesia Indah merupakan aset negara yang menjadi salah satu pilihan tujuan wisata bagi para wisatawan sehingga diharapkan dapat berkontribusi dalam perekonomian negara. Keputusan pemerintah mengambil kembali pengelolaan Taman Mini Indonesia Indah (TMII) dari Keluarga Cendana di bawah Yayasan Harapan Kita (YHK) menuai reaksi beragam terutama pada platform Twitter. Sementara beberapa pengguna mendukung langkah tersebut, yang lain mengungkapkan kekecewaan dan kekhawatiran atas dampaknya terhadap pariwisata dan lapangan kerja. Banyak juga yang mempertanyakan kemampuan pemerintah mengelola TMII secara efektif dan efisien. Secara keseluruhan, sentimen terhadap keputusan tersebut terbagi, dengan komentar positif yang lebih banyak daripada negatif. Data dalam penelitian ini diambil dengan metode crawling data menggunakan Node.js dan diolah pada software rapidminer dengan algoritma Naïve Bayes. Temuan menunjukkan hasil dengan menggunakan 323 komentar dan diproses menggunakan metode Naïve Bayes dan mencapai akurasi keseluruhan sebesar 82.74% sedangkan untuk nilai AUC sebesar 0.500 bahkan hasil pengujian pengklasifikasian AUC nilai keakuratannya dapat dikategorikan sebagai Excellent Classification.

Kata kunci : Data mining, Analisis Sentimen, Naive Bayes, Taman Mini Indonesia Indah, Twitter

ABSTRACT

Tourism plays an important role in the economic development of a country. Taman Mini Indonesia Indah is a state asset which is one of the choices of tourist destinations for tourists so that it is expected to contribute to the country's economy. The government's decision to take back the management of Taman Mini Indonesia Indah (TMII) from the Cendana Family under the Harapan Kita Foundation (YHK) drew mixed reactions, especially on the Twitter platform. While some users support the move, others express disappointment and concern over the impact on tourism and employment. Many also question the government's ability to manage TMII effectively and efficiently. Overall, sentiment towards the decision was divided, with more positive comments than negative. The data in this study were taken by crawling data using Node.js and processed in rapidminer software with the Naïve Bayes algorithm. The findings show the results using 323 comments and processed using the Naïve Bayes method and achieving an overall accuracy of 82.74% while for an AUC value of 0.500 even the results of the AUC classification test the accuracy value can be categorized as Excellent Classification.

Keyword : Data mining, Sentiment Analysis, Naive Bayes, Taman Mini Indonesia Indah, Twitter

1. PENDAHULUAN

Pariwisata memainkan peran penting dalam kemajuan ekonomi suatu negara. Dengan daya tarik Taman Mini Indonesia Indah yang luar biasa, mulai dari rumah adat, museum dan pendopo wisata alam, serta pertunjukan seni dan budaya Indonesia yang

bisa dilihat merupakan salah satu pilihan destinasi wisata para wisatawan dari turis maupun warga lokal.

Pemerintah memutuskan untuk mengambil kembali hak penguasaan dan pengelolaan Taman Mini Indonesia Indah (TMII) dari Keluarga Cendana di bawah Yayasan Harapan Kita (YHK).

Pengambilalihan tersebut tertuang di dalam Peraturan Presiden No.19/2021 tentang Pengelolaan Taman Mini Indonesia Indah. Pemerintah menyampaikan, pengambilalihan pengelolaan Taman Mini Indonesia Indah dilakukan karena sejumlah alasan. Sekretaris Kemensetneg, Setya Utama mengatakan, langkah ini dilakukan demi pengelolaan Taman Mini Indonesia Indah yang lebih baik. Hal tersebut disampaikan Kepala Kantor Staf Presiden (KSP) Moeldoko dalam konferensi pers di Gedung Bina Graha, Jakarta, Jumat (9/4/2021). Moeldoko mengatakan, salah satu yang jadi pertimbangan pengambilalihan yakni kerugian yang dialami pengelola TMII setiap tahun yang nilainya mencapai Rp 40 miliar hingga Rp 50 miliar. Atas kerugian itu, kata Moeldoko, TMII tidak dapat berkontribusi pada keuangan negara.

Analisis sentimen merupakan teknik untuk mengevaluasi dan mengidentifikasi sentimen dan opini positif dan negatif. Data kumpulan opini yang akan dianalisis adalah informasi tekstual yang dapat digali dari kolom komentar Twitter. Dalam Text mining, suatu kelas bersifat topik yang lebih umum, sehingga diperlukan klasifikasi. Ada beberapa algoritma klasifikasi, salah satunya adalah *Naive Bayes* yang digunakan untuk menghitung probabilitas ketidakpastian data. Oleh karena itu, diperlukan proses untuk memilih fitur yang relevan. Oleh karena itu, perlu dilakukan analisis sentimen yang bertujuan untuk menganalisis opini publik terhadap pengambilalihan kepemilikan Taman Mini Indonesia Indah oleh pemerintah.

Penelitian ini diharapkan dapat menjadi informasi untuk melihat bagaimana perubahan pengelolaan yang dilakukan pemerintah untuk tetap menjalankan wisata Taman Mini Indonesia Indah sebagai aset negara yang memiliki nilai keekonomian, nilai sosial budaya, dan beragam nilai didalamnya.

2. METODOLOGI

Metode yang akan digunakan adalah Metode *Knowledge Discovery in Databases* dianggap efektif dan tepat hingga memprosesnya dalam format yang benar untuk tujuan pengambilan keputusan.

Gambar 1. Metode Penelitian



Berikut tahap penelitian menggunakan metodologi *Knowledge Discovery in Databases* (KDD) terdiri dari:

1. Selection

Data selection merupakan tahapan pertama dalam KDD, pada tahap ini akan dilakukan pengumpulan, seleksi, dan pelabelan data. Data dikumpulkan dari sosial media twitter. Kemudian dilakukan proses pelabelan untuk menentukan kalimat bernilai positif dan negatif.

2. Preprocessing

Tahap preprocessing bertujuan untuk mengolah data mentah yang telah dikumpulkan yaitu menghilangkan *noisy* dan *missing value* sehingga dapat digunakan pada tahap berikutnya. Berikut tahap *preprocessing* yang digunakan yaitu:

1) Cleaning

Cleaning dilakukan untuk membersihkan teks dengan tujuan menghapus tanda baca seperti titik, koma, dan lain-lain, serta angka dalam tweet dan elemen lain seperti *HTML*, *URL*, dan *hashtag*.

2) Case Folding

Case folding adalah tahap untuk mengubah bentuk huruf agar seragam. Pada penelitian ini, *case folding* dilakukan dengan mengubah semua huruf menjadi huruf kecil (*lowercase*).

3) Stemming

Proses menemukan kata dasar (*root word*) dari kata berimbuhan dengan cara menghilangkan semua imbuhan yang terdiri dari awalan, sisipan, akhiran dan kombinasi awalan dan akhiran. Dalam *text analytics*,

stemming merupakan salah satu proses penting yang sangat mempengaruhi kualitas hasil analisis.

4) *Tokenizing*

Tokenisasi adalah proses memisahkan kalimat menjadi kata-kata individual untuk setiap kata. Hal ini dilakukan untuk mendapatkan penggalan kata yang akan dijadikan entitas dengan nilai matriks dokumen yang akan dianalisis.

5) *Stopwords*

Tahap filtering dilakukan untuk mengurangi pemilihan kata dalam dokumen teks atau dimensi kata dalam corpus dengan menghilangkan kata-kata yang sering muncul namun tidak informatif, yang disebut sebagai *stopwords*.

3. Data Mining

Data mining atau penambangan data adalah proses analisis data pada kumpulan data besar untuk mengidentifikasi hubungan antar data dan merangkumnya menjadi bentuk yang lebih mudah dipahami dan bermanfaat. Tujuannya adalah untuk menemukan pengetahuan baru yang tersembunyi dalam database dengan menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengidentifikasi informasi yang relevan. Pada proses data mining yaitu mengimplementasikan algoritma *Naive Bayes* tersebut dilakukan dengan menggunakan *Rapidminer*.

4. Evaluation

Untuk mengevaluasi dan memvalidasi hasil penelitian ini, peneliti menggunakan *rapidminer*, data training yang digunakan adalah komentar pada twitter dengan keyword “pengambil alihan tmii” “Taman Mini Indonesia Indah” serta “pemerintah” kemudian dataset dikelompokkan menjadi 2 bagian yaitu positif dan negatif. Lalu untuk memvalidasi data menggunakan *cross validation*. Akurasi diukur dengan menggunakan confusion matrix yang digunakan untuk menghitung nilai *AUC*. Klasifikasi nilai *AUC* yang digunakan adalah sebagai berikut (Gorunescu, 2011):

0,90-100 = *Excellent Classification*;

0,80-0,90= *Good Classification*;

0,70-0,80= *Fair Classification*;

0,60-0,70= *Poor Classification*;

0,50-0,60= *Failure*;

3. LANDASAN TEORI

Analisis Sentimen

Analisis sentimen adalah teknik pengolahan bahasa alami yang digunakan untuk mengekstrak dan menganalisis opini atau sentimen dari teks. Tujuan dari analisis sentimen adalah untuk mengetahui pandangan atau opini yang terkandung dalam teks, apakah positif, negatif, atau netral. Analisis sentimen dapat dilakukan pada berbagai jenis teks, seperti ulasan produk, tweet, atau postingan di media sosial.

Data Mining

Data mining adalah proses ekstraksi informasi yang berguna dari data yang besar, kompleks, dan beragam. Tujuan dari data mining adalah untuk menemukan pola atau hubungan yang tersembunyi dalam data yang dapat digunakan untuk membuat keputusan yang lebih baik. Data mining dapat digunakan pada berbagai jenis data, seperti data transaksi, data pelanggan, atau data sosial.

Naive Bayes

Naive Bayes adalah algoritma klasifikasi yang berdasarkan pada teorema *Bayes*. Algoritma ini digunakan untuk memprediksi kelas atau label dari suatu data berdasarkan pada fitur atau atribut yang dimilikinya. *Naive Bayes* dianggap sebagai algoritma yang sederhana dan efektif dalam melakukan klasifikasi pada data yang besar dan kompleks.

Cross Validation

Cross validation adalah teknik evaluasi model yang digunakan untuk mengukur kinerja suatu model pada data yang tidak digunakan dalam proses pelatihan. Teknik ini dilakukan dengan membagi data menjadi beberapa bagian, di mana satu bagian digunakan sebagai data uji dan bagian lainnya digunakan sebagai data pelatihan. Teknik *cross validation* digunakan untuk menghindari *overfitting* dan memastikan bahwa model yang dibuat dapat digeneralisasi dengan baik pada data yang baru.

4. HASIL DAN PEMBAHASAN

Hasil dari penelitian yang dilakukan adalah sentimen analisis opini publik terhadap pengambil alihan TMII oleh pemerintah

dengan metode *Naïve Bayes*. berikut ini pembahasan mengenai tahapan penelitian yang telah dilakukan:

1. Pengolahan Data

Mengumpulkan data tweet yang diperoleh dengan menggunakan *softwatre Node.js* peneliti mengumpulkan 323 data opini setelah itu dikelompokan menjadi positif dan negatif kemudian lanjut proses pelabelan. Hasil *crawling data* dan pelabelan seperti pada tabel 1.

No	Text	Label
1	belasan ribu pengunjung padati objek wisata taman mini indonesia indah kelurahan ceger kecamatan cipayung jakarta timur saat libur hari lahir pancasila	positif
2	dapat layanan pajak dalam indonesia maju expo dan forum yang selenggara	positif
3	taman mini kita lihat apa gitu saja	negatif
4	taman mini ragunan ancol sudah lebih dari cukup bahagia asal bersama orang kita cinta	positif
5	ambil alih kelola taman mini indonesia indah lakukan karena jumlah alasan sekretaris kemensetneg setya utama katakan langkah ini lakukan demi kelola taman mini indonesia lebih baik	positif
6	dalam waktu bulan pemerintah akan tunjuk mitra baru untuk kelola taman mini indonesa indah kemensetneg akan tunjuk ganti yayasan harapan kita kelola taman mini lama tahun	positif
7	presiden joko widodo terbit peraturan presiden nomor tahun tentang taman mini indonesia indah dalam atur kuasa dan kelola taman mini indonesia indah lakukan kemensetneg	positif

8	selama tahun taman mini indonesia indah kelola yayasan milik keluarga soeharto lama era reformasi nyata banyak aset negara masih belum kelola oleh negara pak jokowi tidak semua pemimpin pikir kembali kaya negara	positif
9	semoga lebih baik lagi setelah ambil alih oleh negara	positif
10	jika pegang pemerintah makin ribet karena dana sumber dari apbd pasti pentok segala aturan lebih baik jika tetap kelola swasta dengan sistem bagi hasil	negatif

Tabel 1. Labeling Data

2. Preprocessing

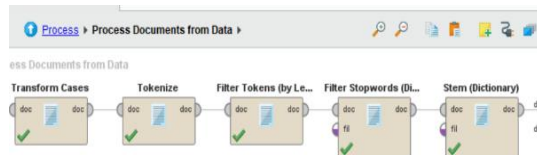
Pada tahap ini dilakukan untuk mengubah data agar sesuai dan dapat diolah pada tahapan selanjutnya. Contoh hasil dari tahap preprocessing yang telah dilakukan dapat dilihat pada tabel 2.

PROSES	HASIL
DATA ULASAN	Kl dibandingkan.. Manajemen yang lama lebih bagus drpd yang skr. Mksd hati pengen untung ternyata buntung. Begitulah nafsu yg gak dibarengin dgn kemampuan.. yg penting pencitraan udh bs ambil alih pengelolaan TMII
CLEANING	Kl dibandingkan Manajemen yang lama lebih bagus drpd yang skr Mksd hati pengen untung ternyata buntung Begitulah nafsu yg gak dibarengin dgn kemampuan yg penting pencitraan udh bs ambil alih pengelolaan TMII
CASE FOLDING	kalau dibandingkan manajemen yang lama lebih bagus daripada yang sekarang maksud hati pengen untung

	ternyata buntung begitulah nafsu yang tidak dibarengin dengan kemampuan yang penting pencitraan sudah bisa ambil alih pengelolaan tmii
STEMMING	jika banding manajemen yang lama lebih bagus daripada sekarang maksud hati ingin untung ternyata buntung begitulah nafsu yang tidak dibarengin dengan kemampuan yang penting pencitraan sudah bisa ambil alih kelola tmii
TOKENIZING	"jika" "banding" "manajemen" "yang" "lama" "lebih" "bagus" "daripada" "yang" "sekarang" "maksud" "hati" "ingin" "untung" "ternyata" "buntung" "begitu" "nafsu" "yang" "tidak" "bersama" "dengan" "kemampuan" "yang" "penting" "citra" "sudah" "bisa" "ambil" "alih" "pengelolaan" "tmii"
STOPWORDS	dibandingkan manajemen lama bagus sekarang maksud hati pengen untung buntung nafsu dibarengin kemampuan penting pencitraan bisa ambil alih pengelolaan tmii

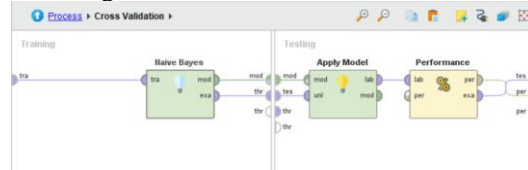
Tabel 2. Hasil Preprocessing

Berikut gambar proses pengolahan data pada rapidminer.



Gambar 2. Proses Pengolahan Data

Berikut gambar proses pengolahan data dengan Cross Validation.



Gambar 3. Proses Cross Validation

3. Evaluasi

Penelitian ini menggunakan algoritma *Naive Bayes* untuk melakukan analisis sentimen. Kemudian dalam menentukan evaluasinya peneliti menggunakan *Accuracy* dan *AUC (Area Under Curve)*.

Pada tahap pengujian metode, peneliti menggunakan *Rapid Miner* dalam melakukan proses eksperimen, *data training* yang digunakan adalah komentar postingan twitter mengenai pengambilalihan Taman Mini Indonesia Indah oleh pemerintah. Kemudian dataset dikelompokkan menjadi 2 (dua) bagian yaitu opini positif dan negatif.

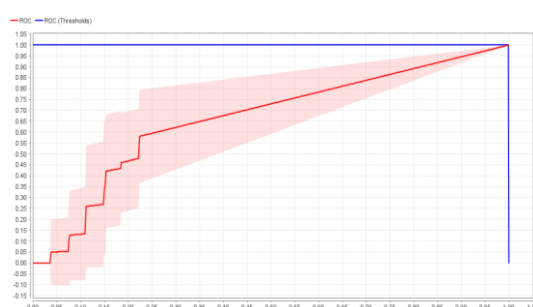
Dari tahapan pengujian yang sudah dilakukan pada dataset opini pengguna Twitter mengenai pengambilalihan Taman Mini Indonesia Indah oleh pemerintah, maka hasil akurasi algoritma *Naive Bayes* sebesar 82.74% sedangkan untuk nilai *AUC* sebesar 0.500. berikut adalah tabel *confusion matrix* algoritma *naive bayes*.

	True Positive	True Negative	
Pred Positive	233	19	92.46%
Pred Negative	34	21	38.18%
	87.27%	52.50%	

Tabel 3. Confusion Matrix

Pada tabel 3 *Confusion Matrix* dapat dilihat, sebanyak 233 data diprediksi *class positive* ternyata sesuai, yaitu masuk kedalam *class positive*, sebanyak 34 data yang diprediksi *class negative* ternyata termasuk kedalam prediksi *class positive*.

Dan sebanyak 19 data yang diprediksi *class positive* ternyata masuk dalam *class negative*, kemudian 21 data di prediksi *class negative* sesuai yaitu termasuk kedalam prediksi *class negative*. *AUC: 0.500 +/- 0.000 (micro average: 0.500) (positive class: negatif)*



Gambar 3. Kurva AUC

5. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, diperoleh bahwa penelitian tentang analisa sentimen opini publik terhadap pengambil alihan TMII oleh pemerintah dengan menggunakan metode *Naïve Bayes* dengan kesimpulan sebagai berikut:

1. Penelitian dilakukan dengan menggunakan 323 komentar dan diproses menggunakan metode *Naïve Bayes* dan mencapai akurasi keseluruhan sebesar 82.74% sedangkan untuk nilai *AUC* sebesar 0.500 bahkan hasil pengujian pengklasifikasian *AUC* nilai keakuratannya dapat dikategorikan sebagai *Excellent Classification*.
2. Hasil analisis sentimen menunjukkan nilai opini positif menempati peringkat pertama dengan jumlah sentimen sebanyak 267. Hasil ini menunjukkan bahwa antusias masyarakat tinggi dalam menanggapi pengambilalihan Taman Mini Indonesia Indah oleh Pemerintah karena dianggap pemerintah dapat mengelola dengan baik manajemen serta seluruh sarana rekreasi yang tersedia sehingga dapat menutup hutang yang ada sebelum diambilalih.

DAFTAR PUSTAKA

- Gorunescu, F. (2011). *Data Mining Concepts, models and Techniques*. Springer Berlin Heidelberg.
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
- Han, J., Kamber, M., & Pei, J. (2011). *Data mining: concepts and techniques*. Elsevier.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
- McCallum, A., & Nigam, K. (1998). A comparison of event models for naive bayes text classification. In *AAAI-98 workshop on learning for text categorization (Vol. 752, pp. 41-48)*.
- Kelleher, J. D., & Tierney, B. (2018). *Data science: an introduction*. CRC Press.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer.
- Romadloni, N. T., Santoso, I., & Budilaksono, S. (2019). Perbandingan Metode Naïve Bayes, KNN dan Decision Tree Terhadap Analisis Sentimen Transportasi KRL Commuter Line. *ikraith-informatika*, 3(2), 1-9.
- Setiawan, K., Supriyadin, I. S., & Buana, R. (2018). Menghitung Rute Terpendek Menggunakan Algoritma A* Dengan Fungsi Euclidean Distance. In *Seminar Nasional Teknologi Informasi dan Komunikasi (pp. 70-79)*.
- Santoso, I., Gata, W., & Paryanti, A. B. (2019). Penggunaan Feature Selection di Algoritma Support Vector Machine untuk Sentimen Analisis Komisi Pemilihan Umum. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 3(3), 364-370.
- Susanti, N. D. (2016). *Uji Perbandingan Akurasi Analisis Sentimen Pariwisata Menggunakan Algoritma Support Vektor Machine dan Naive Bayes (Doctoral dissertation, Magister Sistem Informasi Program Pascasarjana FTI-UKSW)*.
- Mukhtar, H., Gunawan, R., Hariyanto, A., & Mulyana, W. (2022). *Peramalan Kedatangan Wisatawan ke Suatu Negara*

- Menggunakan Metode Support Vector Machine (SVM). *Jurnal CoSciTech (Computer Science and Information Technology)*, 3(3), 274-282.
- Hayuningtyas, R. Y., & Sari, R. (2019). Analisis sentimen opini publik bahasa indonesia terhadap wisata tmii menggunakan naïve bayes dan ps. *Techno Nusa Mandiri*, 16(1), 37-42.
- Rahma, N. A., Garno, G., & Sulistiyowati, N. (2022). Analisis Sentimen Tempat Wisata Di Jakarta Pasca Covid-19 Dengan Algoritma Naïve Bayes. *Jurnal Pendidikan dan Konseling (JPDK)*, 4(6), 5894-5908.
- Sari, R., & Hayuningtyas, R. Y. (2019). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Pada Wisata TMII Berbasis Website. *Indonesian Journal on Software Engineering (IJSE)*, 5(2), 51-60.
- Somantri, O., & Dairoh, D. (2019). Analisis Sentimen Penilaian Tempat Tujuan Wisata Kota Tegal Berbasis Text Mining. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, 5(2), 191-196.
- Fanissa, S., Fauzi, M. A., & Adinugroho, S. (2018). Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(8), 2766-2770.
- Hayuningtyas, R. Y., & Sari, R. (2019). Analisis sentimen opini publik bahasa indonesia terhadap wisata tmii menggunakan naïve bayes dan ps. *Techno Nusa Mandiri*, 16(1), 37-42.
- Direktorat Jenderal Kekayaan Negara. (2021). Pengambilalihan TMII untuk dorong penerimaan negara. <https://www.djkn.kemenkeu.go.id/berita-media/baca/13124/Pengambilalihan-TMII-untuk-dorong-penerimaan-negara.html#:~:text=Pemerintah%20memutuskan%20untuk%20mengambil%20kembali,19%2F2021%20tentang%20Pengelolaan%20TMII> diakses pada 11 Juni 2023, pukul 19.45.
- Sari, N. (2021, April 10). Alasan TMII Diambil Alih Pemerintah: Terus Rugi Puluhan Miliar Rupiah dan Pengelolaan Perlu Dibenahi. *KOMPAS.com*; [04/10/07113061/alasan-tmii-diambil-alih-pemerintah-terus-rugi-puluhan-miliar-rupiah-dan](https://nasional.kompas.com/read/2021/04/10/07113061/alasan-tmii-diambil-alih-pemerintah-terus-rugi-puluhan-miliar-rupiah-dan) diakses pada 11 Juni 2023, pukul 20.00.