

Analisa Sentimen Tentang Ibu Kota Nusantara (IKN) Dengan Menggunakan Algoritma *K-Nearest Neighbors* (KNN) dan *Naïve Bayes*

Choirul Huda¹, Mesra Betty Yel²

^{1,2}Teknik Informatika, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Jakarta, Indonesia
Chrlhuda25@gmail.com¹, optime.mby@gmail.com²

Abstrak– Pemindahan ibu kota bukan merupakan suatu fenomena yang baru. Senin, 17 Januari 2022 nama “Nusantara” diumumkan pertama kali oleh kepala Bappenas Suharso Monoarfa. Panja RUU IKN menyetujui “Nusantara” sebagai nama ibu kota negara setelah delapan fraksi menyetujui. Terletak di Kalimantan Timur, Nusantara adalah sebuah kata yang akan mengganti posisi Jakarta sebagai ibu kota negara Indonesia. Pemindahan Ibu Kota Nusantara (IKN) menimbulkan berbagai macam pro dan kontra. Mulai dari pemilihan lokasi, pengesahan Undang-Undang yang dinilai terlalu terburu-buru, dan akhir-akhir ini pemerintahan juga mengajak masyarakat Indonesia untuk membantu mendukung untuk membangun Ibu Kota Nusantara. Menurut seorang pejabat dari Kementerian Perencanaan Pembangunan Nasional Indonesia, pemerintah bertekad untuk memindahkan ibu kota Indonesia keluar dari pulau Jawa. Pada April 2019, rencana 10 tahun untuk memindahkan semua kantor pemerintahan ke ibukota baru diumumkan Kementerian Perencanaan Pembangunan Nasional merekomendasikan tiga provinsi yaitu Kalimantan Selatan, Kalimantan Tengah, dan Kalimantan Timur yang memenuhi syarat ibu kota baru, termasuk bebas dari gempa bumi, tsunami dan gunung berapi. Dari nama ini pemerintah dan DPR berharap agar ibu kota yang berlokasi di Sepakunegara (sekitar Penajam Paser Utara dan Kutai Kartanegara) menjadi symbol identitas nasional, kota dunia untuk semua dan menjadi penggerak ekonomi Indonesia masa depan.

Kata Kunci: Data Mining, Twitter, Ibu Kota Nusantara, *Naïve Bayes*, *K-Nearest Neighbor*

Abstract– *Moving the capital city is not a new phenomenon. Monday, January 17 2022, the name "Nusantara" was first announced by the head of Bappenas, Suharso Monoarfa. The IKN Bill Working Committee approved "Nusantara" as the name of the country's capital after eight factions agreed. Located in East Kalimantan, Nusantara is a word that will replace Jakarta as the capital of Indonesia. The relocation of the Indonesian Capital City (IKN) raises various pros and cons. Starting from the choice of location, the ratification of the law which was considered too hasty, and recently the government has also invited the Indonesian people to help support building the capital city of the archipelago. According to an official from Indonesia's Ministry of National Development Planning, the government is determined to move Indonesia's capital out of Java. In April 2019, a 10-year plan to move all government offices to the new capital was announced. The Ministry of National Development Planning recommended that three provinces, namely South Kalimantan, Central Kalimantan and East Kalimantan meet the requirements for a new capital city, including being free from earthquakes, tsunamis and volcanoes. From this name, the government and the DPR hope that the capital city located in Sepakunegara (around North Penajam Paser and Kutai Kartanegara) will become a symbol of national identity, a world city for all and a driver of Indonesia's future economy.*

Keywords: Data Mining, Twitter, Capital City of the Archipelago, *Naïve Bayes*, *K-Nearest Neighbor*.

1. PENDAHULUAN

Pemindahan ibu kota bukan merupakan suatu fenomena yang baru[1]. Pemindahan ibu kota Indonesia dari Jakarta ke daerah lainnya sudah sering dibahas sejak kepresidenan era Soekarno sampai era Susilo Bambang Yudhoyono namun masih belum bisaterlaksana. Melalui rapat yang dilaksanakan pada tanggal 29 April 2019, Presiden RI JokoWidodo menetapkan untuk memindahkan ibu kota ke luar pulau Jawa.

Pemindaan Ibu Kota Nusantara (IKN) menimbulkan berbagai macam pro dan kontra. Mulai dari pemilihan lokasi, pengesahan Undang-Undang yang dinilai terlalu terburu-buru, dan akhir-akhir ini pemerintahan juga mengajak masyarakat Indonesia membantu mendukung untuk membangun Ibu Kota Nusantara. Pemerintahan presiden Jokowi Widodo mempertimbangkan untuk memindahkan ibukota dari Jakarta, dengan rencana untuk menyelesaikan penilaian situs alternatif potensial untuk ibu kota baru Indonesia padaakhir 2017. Pada April 2019, rencana 10 tahun untuk memindahkan semua kantor pemerintahan ke ibukota baru diumumkanKementerian Perencanaan Pembangunan Nasional merekomendasikan tiga provinsi yaitu Kalimantan Selatan, Kalimantan Tengah, dan Kalimantan Timur yangmemenuhi syarat ibu kota baru, termasukbebas dari gempa bumi, tsunami dan gunungberapi.

Ibu Kota Negara (IKN) baru yang dinamakanNusantara secara resmi dipakai sejak UU IKN No. 3 tahun 2022 ditandatangani Presiden Jokowi pada tanggal 15 Februari 2022. Dari nama ini pemerintah dan DPR berharap agar ibu kota yang berlokasi di Sepakunegara (sekitar Penajam Paser Utara dan Kutai Kartanegara) menjadi symbol identitas nasional, kota dunia untuk semua dan menjadi penggerak ekonomi Indonesia masa depan.

2. METODE PENELITIAN

a. Data Penelitian

Data yang digunakan dalam penelitian merupakan data public. Data mentah yang diperoleh berjumlah 5000 record dengan 3 atribut dapat dilihat pada tabel 3.1 sebagai berikut:

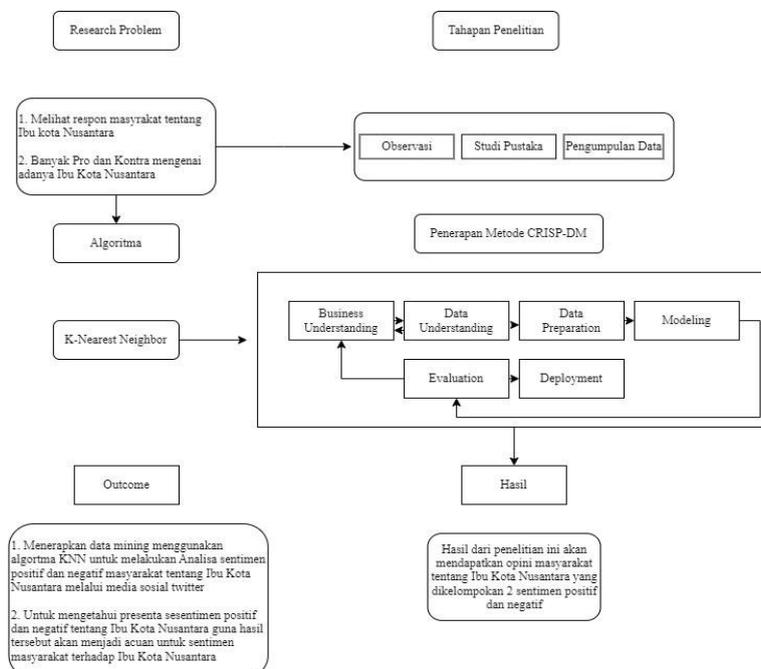
Tabel 1. Data mentah

Atribut	Keterangan
Tweet	Komentar masyarakat melalui media social twitter
Negative	Sentiment buruk atau negative
Positive	Sentiment baik atau positive

Data tersebut didapatkan dari twitter menggunakan ibu kota nusantara, setelah tweet di temukan diperlukan API twitter untuk mendapatkan data tweets tersebut , lalu proses selanjut nya dibutuhkan aplikasi rapidminer untuk mendapatkan data tweet menggunakan extention API twitter yang ada di dalam aplikasi rapidminer ,proses selanjutnya yang dilakukan pendeskripsian data, pemilihanatribut dan cleansing data. Setelah itu data di *export* menjadi data csv agar memudahkan proses Analisa. Pada tahap pengumpulan data ini yang di kumpulkan merupakan data tweet dari tanggal 15 Februari 2022 sampai 15 Maret2023. Data yang digunakan dalam penelitian merupakan data kualitatif. Data kualitatif adalah data yang mencirikan atau menggambarkan sesuatu[2], data yang dapat diamati dan direkam. Data kualitatif dibagi menjadi 2 yaitu Data Primer dan Data Sekunder.

b. Penerapan Metodologi

Pada tahapan penelitian ini akan dilakukan pengumpulan data awal maka proses selanjutnya akan dilakukan penarikan data, kemudian setelah itu dilakukan Cleansing Data, lalu pada tahap selanjutnya dilakukan Pemilihan Atribut, tahap berikutnya melakukan penerapan metode CRISP-DM yang pada proses penelitian ini mengacu pada enam tahap CRISPM-DM.



Gambar 1. Tahapan Penerapan Metodologi

c. Rancangan Pengujian

Pada penelitian ini metode yang digunakan adalah *K-Nearest Neighbors Algorithm Classification*. Data yang digunakan ialah tweet yang diambil dari aplikasimedia sosial twitter menggunakan API Twitterdengan cara menggunakan ekstensi

dari Rapidminer yaitu *search twitter*[3] yang di ambil dari 15 Februari2022 sampai 15 Maret 2023 setelah data tersebut di dapatkan lalu masuk ke proses penerapan CRISP-DM yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi dan penyebaran (*Deployment*). Data tersebut akan diproses dalam pengujian algoritma *K-Nearest Neighbors* menggunakan tools Rapidminer supaya dapat nilai akurasi klasifikasi yang baik dan optimal.

Pada tahapan ini pendekatan yang digunakan menggunakan metode Cross Industry Standard for Data Mining (CRISP-DM). Proses penelitian ini mengacu pada enam tahap terdapat dalam CRISP- DM ini yakni dijelaskan sebagai berikut:

1. **Pemahaman Bisnis (Business Understanding)**

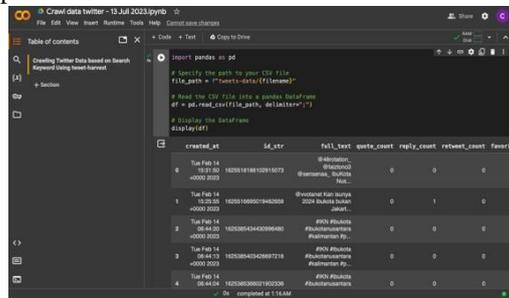
Ibu Kota Nusantara adalah ibu kota masa depan Indonesiayang rencananya akan diresmikan pada 17 Agustus 2024, bersamaan dengan perayaan hari kemerdekaan Indonesia ke-79. Ibu Kota Nusantara direncanakan akan menggantikan Jakarta yang telah menjadi ibu kota sejak 1961. Terletak dipantai timur pulau Kalimantan.

2. **Pemahaman Data (Data Understanding)**

Memahami informasi data yang akan digunakan dalam penelitian, adapun beberapa tahap untuk melakukan penelitian ini yaitu:

a) **Pengumpulan Data**

Data yang dibutuhkan untuk mendukung dalam melakukan pemahaman data[4]. Adapun sumber data utama yang digunakan dalam penelitian ini adalah data tweet ibu kotanusantara pada 15 Februari 2022 sampai 15



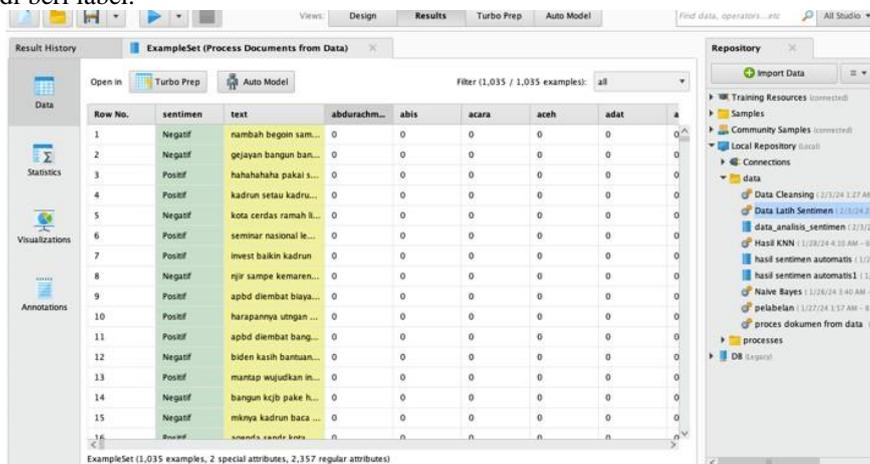
Gambar 2. Model Proses Pengumpulan Data Maret 2023

b) **Cleansing Data**

Pada tahapan ini peneliti akan melakukan proses Cleansing dengan menggunakan *gataframework.com* sebelum melakukan pemberian label pada sentimen, agar dataset pada file xls terbaru terhindar dari duplikasi datadan tanda yang tidak diperlukan. Pada proses iniakan dilakukan pembersihan dari berbagai noiseseperti menghilangkan link URL, username, retweet, digit angka dan karakter. Setelah selesai menjalankan proses Cleansing denganmenggunakan *gataframework.com*.

c) **Labeling Data**

Melakukan klasifikasi atau penentuan sentimenberdasarkan masing-masing tweet. Selama proses penentuan sentimen berlangsung, disini peneliti melakukan pemberian sentimen secara manual[5] . Berikut ini data yang sudah di beri label.

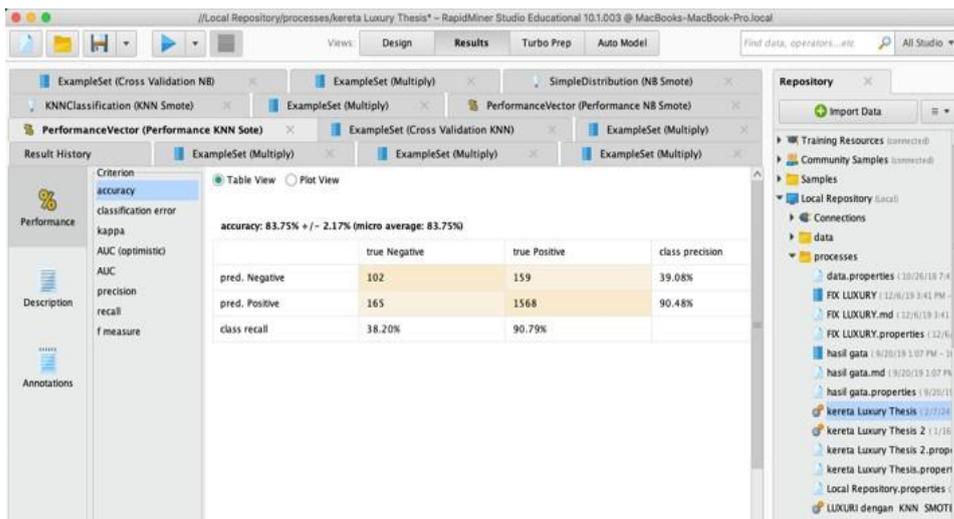


Gambar 3. Hasil Penentuan Sentimen

3. HASIL DAN PEMBAHASAN

Pada penelitian ini akan dilakukan pengujian menggunakan *confusion matrix* untuk melihat hasil pengujian data yang diperoleh dari tahapan modelling dengan menggunakan algoritma K-NN. Total dataset yang dikumpulkan adalah 2.006 data. Berikut ini adalah *confusion matrix* hasil dari tahapan modelling yang dapat dilihat hasil dari perhitungan hasil RapidMiner dapat membuktikan hasil prediksi dari model K-NN.

Gambar 4. Hasil Cleansing di Gataframework



Gambar 5. Hasil Accuracy Confusion Matrix

Berdasarkan pada Gambar 3.1 dapat dibuat kesimpulan bahwa *precision* menunjukkan tingkat ketepatan data yang diprediksi positif terhadap banyaknya data yang benar diprediksi negatif yang menghasilkan presentase ketepatan adalah 39,08 %, sedangkan untuk data bersentimen positif memiliki *precision* sebesar 90,48 %. Untuk sentimen negatif memiliki *recall* (*Specificity*) adalah 83,75 % sehingga dapat disimpulkan bahwa model dapat menemukan kembali informasi atau data yang benar-benar negatif dengan sangat baik, sedangkan pada sentimen positif memiliki *recall* sebesar 90,79 %. Nilai *accuracy* yang dihasilkan menggunakan model *K-Nearest Neighbor* adalah 83,75 %, sehingga dapat disimpulkan bahwa algoritma *K-Nearest Neighbor* dapat mengklasifikasi sentimen dengan baik menggunakan data Ibu Kota Nusantara.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan mengenai klasifikasi opini masyarakat Indonesia terhadap Ibu Kota Nusantara Indonesia dapat diambil kesimpulan yaitu hasil Sentimen masyarakat terhadap Ibu Kota Nusantara Indonesia ditemukan 87% bersentimen positif dan 13% bersentimen negatif. Maka dari hasil tersebut dapat ditarik kesimpulan bahwa masyarakat Indonesia sebagian besar memiliki respon negatif terhadap Piala Dunia u20 Indonesia dan berdasarkan klasifikasi model algoritma *K- Nearest Neighbor*, menggunakan split data perbandingan 0.8 : 0.2 dengan nilai k = 3

terhadap dataset Ibu Kota Nusantara didapatkan nilai accuracy sebesar 83.75%. Sehingga dapat dikatakan bahwa algoritma K-KNN dapat mengklasifikasikan data secara baik dan benar.

REFERENSI

- [1] D. Aryanti, “Analisis Sentimen Ibukota Negara Baru Menggunakan Metode Naïve Bayes Classifier,” *J. Inf.Syst. Res.*, vol. 3, no. 4, pp. 524–531, 2022, doi: 10.47065/josh.v3i4.1944.
- [2] M. T. Student *et al.*, “No 主観的健康感を中心とした在宅高齢者における健康関連指標に関する共分散構造分析Title,” *Front. Neurosci.*, vol. 14, no. 1, pp. 1–13, 2021.
- [3] D. G. Nugroho, Y. H. Chrisnanto, and A. Wahana, “Analisis Sentimen Pada Jasa Ojek Online ... (Nugroho dkk.),” pp. 156–161, 2015.
- [4] F. Nurhuda, S. W. Sihwi, and A. Doewes, “Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier,” *ITSmart J. Teknol. dan Inf.*, vol. 2, no.2, pp. 35–42, 2013.
- [5] N. L. P. M. Putu, Ahmad Zuli Amrullah, and Ismarmiaty, “Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation,” *J. RESTI (Rekayasa Sist. dan Teknol.Informasi)*, vol. 5, no. 1, pp. 123–131, 2021, doi: 10.29207/resti.v5i1.2587.