

SENTIMENT ANALYSIS PUBLIC OPINION OF CFW (CITAYAM FASHION WEEK) ON MEDIA SOCIAL TWITTER USING NAÏVE BAYES CLASSIFIER

Angga Aditya Permana¹⁾, Permana Perdana Putra²⁾,

¹ Program Studi Informatika, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara, Kabupaten Tangerang, Banten

² Program Studi Teknik Informatika Fakultas Teknik, Universitas Muhammadiyah Tangerang, Kota Tangerang, Banten

Co Responden Email: angga.permana@umn.ac.id

Abstract

Riwayat

Received 13 Dec 2022

Revised 10 Jan 2023

Accepted 01 Feb 2023

Available online 15 Feb 2023

Keywords:

Sentiment Analysis,
Citayam,
Fashion Week,
Naïve Bayes Classifier

The increasing interest of young people in Indonesia towards local fashion brands is an opportunity to bring the Indonesian fashion industry to the international market, this has triggered people's desire to become a fashion designer. This fashion development has made many teenagers able to express their hobbies, one of which is on public roads in the Sudirman area, or more commonly called Citayam Fashion Week (CFW). CFW is very much a phenomenon in the middle of 2022, but of course this reaps a lot of opinions, some react positively, some even criticize, This is the area where the study needs to be done in order to examine the sentiment on Twitter and other social media, The research methodology consists of several stages, namely, data collection, praprocessing, data classification, and conclusions and suggestions the method used is naïve Bayes classifier with evaluation using a confusion matrix, with an accuracy of 84%. This research is useful for extracting opinions from Twitter social media regarding CFW with the results of more positive responses to CFW.

Abstrak

Riwayat

Diterima 13 Des 2022

Revisi 10 Jan 2023

Disetujui 01 Feb 2023

Terbit online 15 Feb 2023

Kata Kunci :

Analisis Sentimen,
Citayam,
Fashion, Week,
Naïve Bayes Classifier

Meningkatnya minat para remaja di Indonesia terhadap *fashion brand* lokal menjadi sebuah kesempatan untuk membawa *industry fashion* Indonesia kepasar internasional, hal tersebut memicu keinginan masyarakat untuk menjadi seorang *fashion designer*. Perkembangan *fashion* inilah yang membuat banyak para remaja untuk dapat mengekspresikan hobinya, salah satunya yaitu di jalan umum tepatnya di kawasan sudirman, atau lebih sering disebut Citayam Fashion Week (CFW). CFW sangatlah menjadi sebuah fenomena pada pertengahan tahun 2022, tetapi banyak sekali masyarakat memiliki pandangan yang berbeda, ada yang menyikapinya dengan positif, adapula yang malah mengkritik, dari sinilah penelitian ini perlu dilakukan untuk dapat menganalisa sentiment yang ada pada media sosial yaitu twitter, langkah penelitian terbagi menjadi beberapa fase yaitu, pengambilan data, preproccesing, klasifikasi data, dan Kesimpulan serta saran lalu metode yang diimplementasikan yaitu naïve bayes classifier dengan evaluasi menggunakan confusion matrix, dengan hasil akurasi sebesar 84%. Penelitian ini bermanfaat untuk mengekstraksi opini – opini dari media social twitter terkait CFW dengan hasil lebih banyak respon positif terhadap CFW.

PENDAHULUAN

Melihat meningkatnya remaja memakai fashion brand lokal bertujuan untuk membawa industri fashion Indonesia kepasar internasional, membuat meningkatnya keinginan masyarakat untuk masuk ke dunia

model profesi seperti fashion designer menjadi cukup diminati.

Fashion merupakan bidang yang dianggap penting bagi semua masyarakat. Kegemaran berpakaian ini salah satu aspek kebutuhan beberapa orang untuk mengekspresikan identitas mereka. Perkembangan fashion telah

JIKA | 112

memicu beberapa yang cukup besar di sudirman seperti CFW (Citayam Fashion Week), platform umum untuk para industri fashion terkait dengan desain, gaya, dan tren terbaru. CFW (Citayam Fashion Week) ini juga bertujuan untuk membawa remaja mengenal brand lokal yang ada di Indonesia (Fashion et al., 2022).

CFW (Citayam Fashion Week) telah dikomentari pada masyarakat. Data yang ada pada media social dapat digunakan sebagai informasi awal dan dimanfaatkan untuk menentukan keputusan, dalam melakukan penggalian informasi berupa text untuk mendapatkan sumber informasi dari *Tweet* pada *Twitter*. *Text mining* adalah kegiatan menggali informasi dengan sumber utama adalah text atau tulisan (Syadid, 2019). Penelitian ini melakukan text mining (Singh et al., 2021) pada suatu komentar mengenai Citayam Fashion Week pada media sosial twitter yang sering di gunakan untuk bersosialisasi antar pengguna sebagai bentuk luapan opini pada media social. (Steven & Wella, 2020)(Lestari et al., 2017)(Destitus et al., 2020). Karena banyaknya luapan opini pada media social, hal ini dapat dijadikan sebagai acuan dasar dalam menganalisa sentiment pada masing masing luapan opini masyarakat terhadap suatu topik yang saat itu sedang menjadi pembicaraan yang cukup menarik.

METODE PENELITIAN

Pada penelitian ini melakukan empat tahapan penelitian, yaitu :

A. Pengumpulan Data

Tahap pertama yang dilakukan adalah mengumpulkan data dengan banyak cara, seperti metode kepustakaan yaitu dengan mengumpulkan jurnal, literatur, makalah, dan situs internet sebagai sumber pustaka dengan bahan tulisan, seperti penelitian saat ini analisis sentimen dengan metode naïve bayes classifier (NBC) (Zulfa & Winarko, 2017)(Fairuz et al., 2021)(Fauzi & Adinugroho, 2018). Data yang diperoleh merupakan sumber data dari twitter (Syakuro, 2017)(Patil & Kolhe, 2021)(Zuhdi et al., 2019) dengan

menggunakan tools Python dan kata kunci Citayam Fashion Week.

B. Preprocessing Data

Pada tahap kedua ini adalah lanjutan dari tahap pertama, yaitu data yang telah dikumpulkan pada tahap sebelumnya, tahap *preprocessing* ini dilakukan untuk menghasilkan dataset yang sesuai dengan kebutuhan penggunaan.

C. Klasifikasi Data

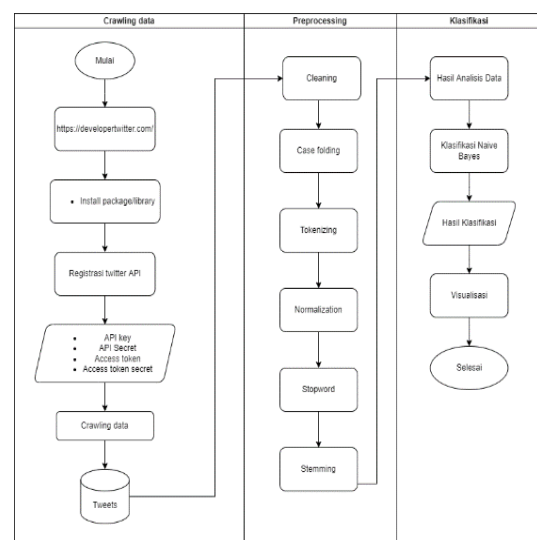
Pada tahap ini, dataset sudah disesuaikan dengan kebutuhan klasifikasi, lalu menggunakan algoritma naïve bayes klasifikasi untuk proses klasifikasi data.

D. Kesimpulan dan saran

Ditahap keempat ini, ditariklah sebuah kesimpulan dari tahap sebelumnya dan memberikan sebuah saran untuk peneliti selanjutnya.

HASIL DAN PEMBAHASAN

Alur Sentimen Analisis pada penelitian ini tergambar pada gambar 1, serta kata kunci yang digunakan adalah Citayam Fashion Week hingga mendapatkan hasil klasifikasi dengan menggunakan Python (Permana et al., 2021)(Alam et al., 2021)(Maddumala et al., 2022) dan algoritma naïve bayes classifier



Gambar 1 Alur Sentimen Analisis

A. Crawling Data

Tahap ini adalah pengambilan data yang telah di lakukan terhadap opini

masyarakat yang berasal dari media sosial twitter. Data yang di ambil berupa tweet dengan teknik crawling sebanyak 500 data tweet, data diambil pada tanggal 30 juli 2022 sebanyak 157 tweet dengan kata kunci “Citayam Fashion Week”. Data didapatkan dengan bantuan menggunakan tools Python JupyterLab dan Twitter API. Sebelum ke tahap pengambilan data kita harus mempunyai API Key untuk mengidentifikasi dengan twitter.

Langkah untuk mendapatkan API key :

1. Login akun Twitter.
2. Masuk ke App Twitter atau buka <https://apps.twitter.com/app/new>
3. Lalu Sign In API Development for Application on Twitter
4. Lalu menginput form dengan informasi yang dibutuhkan.

Selanjutnya pada tahap crawling data, yang dapat dilakukan adalah autentifikasi python dengan Twitter, selanjutnya bisa dilakukan pencarian tweet dengan katakunci yang diinginkan, setelah mendapatkan sejumlah tweet pada tahap crawling data, bisa langsung di simpan dengan format csv untuk dapat dilanjutkan ketahap preprocessing.

B. Pre-processing Data

Langkah kedua ini dilakukan untuk membersihkan dataset dari kata yang tidak diperlukan serta kata yang tidak memiliki makna. Ada beberapa tahapan dalam preprocessing yaitu *Case Folding*, *Cleansing*, *Tokenizing*, *Normalization*, *Stopword Removal*, dan *Steaming*.

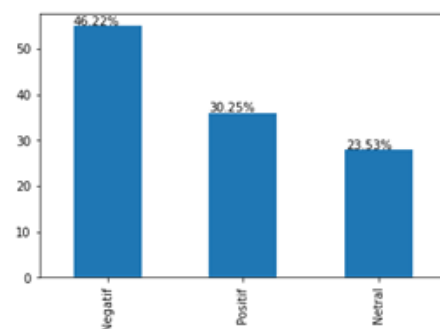
C. Translate

Tahap ini adalah translate, berguna untuk menerjemahkan data tweets dari tweets bahasa indonesia ke bahasa inggris, tahap translate menggunakan library vader sentiment untuk translate ke bahasa inggris. Contoh kata translate dapat dilihat pada table 1.

Tabel 1. Contoh kata translate

Before	After
tren citayam fashion week trik gaya keren anti norak viva citayamfashionweek cfw	the trend of citayam fashion week cool style tricks anti tacky viva citayamfashionweek cfw
tren citayam fashion week trik gaya keren anti norak	Citayam fashion week trend anti-tacky cool style trick

D. Hasil Klasifikasi



Gambar 2. Hasil Klasifikasi

Gambar 2 diatas menunjukkan bahwa, dari total data terdapat sebanyak 46,22% twit yang masuk sebagai kategori negative, 30,25% positif dan 23,53 netral.

Selanjutnya tahap ini adalah mengklasifikasikan data tweet dengan proses labelling sebagai sentimen positif, negatif, atau netral dilakukan untuk menentukan score polarity. Pada tahap ini menggunakan metode Naïve Bayes Classifier

Setelah dilakukan tahapan penerjemahan pada tahapan *preprocessing*, lalu melakukan klasifikasi naïve bayes dengan menggunakan *library python* yaitu *TextBlob*. Untuk label *positif* adalah jika *score* > 0, label *negatif* adalah jika *score* < 0 dan untuk netral adalah jika *score* = 0. Contoh *score* adalah 0,30, *score* 0,30 termasuk kedalam label *positif* karena nilai *score* > 0, contoh *score* adalah -0,30 termasuk kedalam label *negatif* karena nilai *score* < 0, dan contoh *score* adalah 0,0 termasuk kedalam label *netral* karena *score* 0.

Hasil klasifikasi *confusion* dan *naïve bayes* adalah hasil dari data yang telah di lakukan tahap *preprocessing* dan di *translate*

lalu ketahap klasifikasi dengan menggunakan *library python* yaitu *TextBlob* data akurasi yang dihasilkan adalah 84%. Sebagai contoh dibawah ini tabel confusion matrix :

Tabel 2. Evaluasi Model

Kelas	Presiisi	Recall	F1-Score	Akurasi
Positif	0.74	0.94	0.83	0.84
Netral	1.00	0.14	0.25	
Negatif	0.93	0.94	0.93	

Tabel 3. Hasil presiisi, recall, F1-Scor

		Predicted Value		
		Positif	Netral	Negatif
Actual Value	Positif	180	120	15
	Netral	213	148	5
	Negatif	17	115	40



Gambar 3. Hasil Tampilan Wordcloud Negatif



Gambar 4. Hasil Tampilan Wordcloud Positif



Gambar 5. Hasil Tampilan Wordcloud Netral

Dari gambar 3 dapat terlihat dengan jelas beberapa kata yang paling sering digunakan yaitu : fashion week, phenomenon, mmc millenials, citayam fashion, dan abg. Sedangkan pada gambar 4 kata yang paling sering digunakan yaitu : fashion week, citayam fashion, phenomenon, children dan smoking. Lalu pada gambar 5 kata yang paling sering digunakan yaitu : fashion week, citayam fashion, citayamfashionweek. cfw, dan bima arya.

KESIMPULAN

Berdasarkan ujicoba yang dilakukan algoritma Naïve Bayes Classifier yang telah dilakukan ada beberapa hal yang dihasilkan oleh penulis, yaitu :

1. Metode *Naïve Bayes Classifier* terbukti efektif dalam klasifikasi sentimen Citayam Fashion Week. Tetapi dengan ada nya tambahan tools dari Confusion Matrix hasil perhitungan akurasi meningkat hingga 84%
2. Metode Naive Bayes Classifier terbukti dapat mengklasifikasikan sentimen dengan sangat baik terhadap Citayam Fashion Week.

Ucapan Terimakasih

kepada semua pihak yang sangat membantu dalam penelitian ini, Universitas Multimedia Nusantara dan Universitas Muhammadiyah Tangerang serta berbagai pihak sehingga kegiatan ini selesai dilakukan.

REFERENSI

- Alam, M. F., Singh, R., & Katiyar, S. (2021). Customer Segmentation Using K-Means Clustering in Unsupervised Machine Learning. In *Proceedings - 2021 3rd International Conference on Advances in Computing, Communication Control and Networking, ICAC3N 2021* (pp. 94–98). <https://doi.org/10.1109/ICAC3N53548.2021.9725644>
- Destitus, C., Wella, W., & Suryasari, S. (2020). Support Vector Machine VS Information Gain: Analisis Sentimen Cyberbullying di Twitter Indonesia. *Ultima InfoSys: Jurnal Ilmu Sistem Informasi*, 11(2), 107–111.

- <https://doi.org/10.31937/si.v11i2.1740>
- Fairuz, A. L., Ramadhani, R. D., & Tanjung, N. A. F. (2021). Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter. *Journal of Dinda : Data Science, Information Technology, and Data Analytics*, 1(1), 42–51. <https://doi.org/10.20895/dinda.v1i1.180>
- Fashion, C., Fenomena, W., Ekonomi, P., Fashion, C., Fenomena, W., Ekonomi, P., & Pemerintah, K. (2022). *Jurnal Mimbar Muamalah*. August, 8–18.
- Fauzi, M. A., & Adinugroho, S. (2018). Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking Optimasi Sisa Bahan Baku Pada Industri Mebel Menggunakan Algoritma Genetika View project Automatic Essay Scoring View project. August.
- Lestari, A. R. T., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Tentang Opini Pilkada DKI 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naïve Bayes dan Pembobotan Emoji. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(12), 1718–1724.
- Maddumala, V. R., Chaikam, H., Velanati, J. S., Ponnaganti, R., & Enuguri, B. (2022). Customer Segmentation using Machine Learning in Python. In *7th International Conference on Communication and Electronics Systems, ICCES 2022 - Proceedings* (pp. 1268–1273). <https://doi.org/10.1109/ICCES54183.2022.9836018>
- Patil, R. S., & Kolhe, S. R. (2021). Resource Creation for Sentiment Analysis of Under-Resourced Language: Marathi. In *Communications in Computer and Information Science* (Vol. 1380, pp. 445–457). https://doi.org/10.1007/978-981-16-0507-9_37
- Permana, A. A., Fahrezi, M. F., Kristiyanti, D. A., & Sihotang, M. (2021). Sentimen Analisis Opini Masyarakat Pada Media Sosial Twitter Terhadap Vaksin Berbayar Menggunakan Metode Naïve Bayes Classifier (Nbc). *Jurnal Teknik*, 10(2), 113–121. <https://doi.org/10.31000/jt.v10i2.5471>
- Singh, K. N., Dorendro, A., Devi, H. M., & Mahanta, A. K. (2021). Analysis of Changing Trends in Textual Data Representation. In *Communications in Computer and Information Science* (Vol. 1380, pp. 237–251). https://doi.org/10.1007/978-981-16-0507-9_21
- Steven, C., & Wella, W. (2020). The Right Sentiment Analysis Method of Indonesian Tourism in Social Media Twitter. *IJNMT (International Journal of New Media Technology)*, 7(2), 102–110. <https://doi.org/10.31937/ijnmt.v7i2.1732>
- Syadid, F. (2019). Analisis Sentimen Komentar Netizen Terhadap Calon Presiden Indonesia 2019 Dari Twitter Menggunakan Algoritma Term Frequency-Invers Document Frequency (Tf- Idf) Dan Metode Multi Layer Perceptron (Mlp) Neural Network. *Skripsi Universitas Islam Negeri Syarif Hidayatullah Jakarta*, 1–89.
- Syakuro, A. (2017). Pada Media Sosial Menggunakan Metode Naïve Bayes Classifier (NBC) Dengan Seleksi Fitur Information Gain (IG) Halaman Judul Skripsi Oleh : Abdan Syakuro. *Analisis Sentimen Masyarakat Terhadap E-Commerce Pada Media Sosial Menggunakan Metode Naive Bayes Classifier (NBC) Dengan Seleksi Fitur Information Gain (IG)*, 1–89.
- Zuhdi, A. M., Utami, E., & Raharjo, S. (2019). Analisis sentiment twitter terhadap capres Indonesia 2019 dengan metode K-NN. *Jurnal Informa Politeknik Indonusa Surakarta*, 5, 1–7.
- Zulfa, I., & Winarko, E. (2017). Sentimen Analisis Tweet Berbahasa Indonesia Dengan Deep Belief Network. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 11(2), 187. <https://doi.org/10.22146/ijccs.24716>