

Analisis Sentimen Teknologi Ai Terhadap Desainer Grafis Menggunakan Naïve Bayes Dengan Metode Pengujian 10 Fold Cross Validation

Qodri Nurfalah^{1,*}, Fuad Nur Hasan²

¹Informatika/Teknik Informatika , Universitas Bina Sarana Informatika, Jakarta, Indonesia, 16424

²Informatika/Teknik Informatika , Universitas Bina Sarana Informatika, Jakarta, Indonesia, 16424

15200270@bsi.ac.id¹, fuad.fnu@bsi.ac.id²

Abstrak

Penelitian ini di peruntukan untuk mengetahui sentimen para desainer grafis terhadap teknologi yang belum lama ini sedang di perbincangkan yaitu teknologi ai. Dengan menggunakan metode Naive Bayes dengan 10-Fold Cross Validation, penelitian ini meneliti bagaimana teknologi AI berdampak pada desainer grafis. Data dikumpulkan melalui crawling platform YouTube. Data yang di dapatkan dari crawling yaitu 1754 data. Dengan melalui tahap pre-processing yaitu, cleansing, tokenize, stopword, stemming, dan lemmatization Hasilnya menunjukkan 796 sentimen negatif dan 958 sentimen positif. Dengan menggunakan matrix confusion, evaluasi ini menunjukkan akurasi 73%. Akurasi tertinggi adalah di fold ke-2 (76%) dan terendah adalah di fold ke-6 (65%), dengan rata-rata akurasi 10 fold adalah 70%. Hasil penelitian menunjukkan bahwa Naive Bayes dengan Validasi Jalur Sepuluh Potongan dapat digunakan dengan cukup akurat untuk menilai sentimen teknologi AI terhadap desainer grafis. Dengan naïve bayes terbukti cukup bagus untuk efektifitas dalam klasifikasi sentimen komentar para desainer grafis

Kata kunci: Analisis Sentimen, Naive Bayes, 10-Fold Cross Validation, Teknologi AI, Desainer Grafis, YouTube

A. Pendahuluan

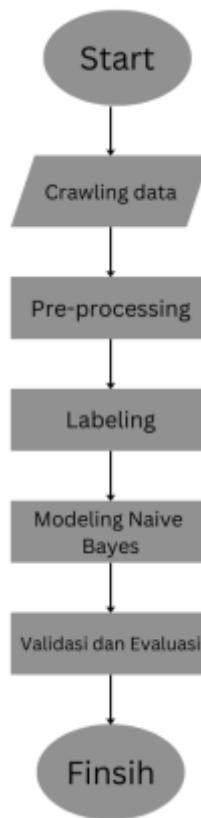
Fokus penelitian ini adalah bagaimana teknologi kecerdasan buatan (AI) telah mengubah banyak aspek kehidupan manusia, mulai dari otomatisasi tugas-tugas sederhana hingga kemampuan untuk menganalisis data yang kompleks. "Kecerdasan buatan, juga dikenal sebagai AI, adalah sistem komputer yang memiliki kemampuan untuk melakukan tugas-tugas yang biasanya membutuhkan kecerdasan manusia"(Sobron et al., 2022). "Teknologi AI yang sangat marak digunakan saat ini telah mendapat banyak perhatian dari masyarakat dan pelaku industri kreatif. Lalu muncul berbagai pro dan kontra hingga pernyataan mengenai pekerjaan ilustrator yang dapat digantikan oleh teknologi AI"(Nur Fadilla & Munadiyah Ramadhani, 2023). Dengan menggunakan algoritma Naive Bayes, penelitian ini menganalisis persepsi desainer

grafis terhadap teknologi kecerdasan buatan. “Salah satu algoritma yang termasuk dalam pendekatan klasifikasi adalah Naive Bayes, sebuah algoritma pengklasifikasian probabilistik sederhana yang menghitung sekumpulan kemungkinan dengan menggabungkan frekuensi dan nilai dari dataset yang diberikan” (Maulidah et al., 2021). Hasil analisis akan diuji menggunakan metode confusion matrix dan validasi menggunakan 10-fold cross-validation untuk mengevaluasi akurasi dan kinerja model yang digunakan. “Pengujian klasifikasi dilakukan untuk 2 mengetahui nilai akurasi algoritma yang akan dievaluasi dan apakah model klasifikasi yang dibuat layak untuk digunakan. Metode evaluasi yang digunakan untuk mengevaluasi kinerja algoritma ini adalah matriks confusion” (Nugroho & Religia, 2021). Dalam hal ini data sentimen di ambil dari platfrom besar yaitu youtube, “Kolom komentar di YouTube adalah tempat di mana orang dapat berbagi pendapat mereka secara bebas, yang menghasilkan banyak data yang penuh dengan informasi tentang perasaan mereka. YouTube juga media sosial yang paling banyak digunakan oleh berbagai demografi” (Utomo et al., 2023).

B. Metode

Naive Bayes

Dalam analisis sentiment ini, memilih menggunakan klasifikasi naïve bayes. Karena algoritma sederhana untuk klasifikasi yang didasarkan pada teorema Bayes dan asumsi independensi yang kuat (naive) antar fitur dan variabel prediktor. “Naïve Bayes Classifier adalah metode klasifikasi yang cepat dan akurat untuk volume data yang besar. Dimana nilai atribut kelas tertentu tidak bergantung pada nilai atribut lainnya” (Kusuma Wardani & Arum Sari, 2021). Untuk analisis sentimen, yang berarti mengkategorikan teks menjadi sentimen positif, negative.



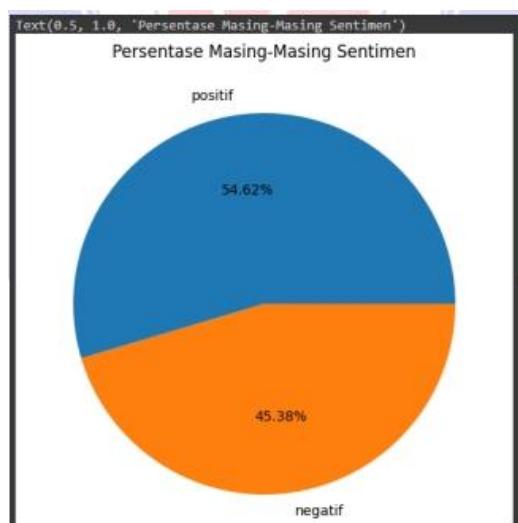
Gambar 1. Urutan flowchart yang digunakan dalam analisis sentimen

Carawling Data

“Crawling adalah teknik pengumpulan data yang digunakan untuk mengindeks informasi pada halaman menggunakan URL dengan menyertakan API”(Miranda et al., 2023). Crawling data, juga dikenal sebagai web crawling atau web scraping, adalah proses mengumpulkan data dari situs web secara otomatis menggunakan program atau bot yang disebut sebagai "crawler" atau "spider". Tujuan utama dari web crawling adalah untuk mengumpulkan informasi tertentu dari halaman web tersebut..

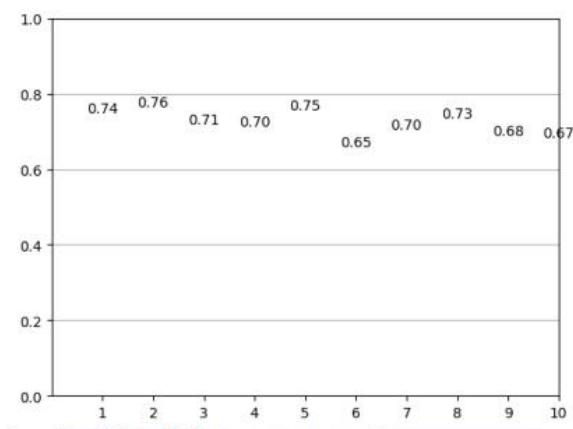
C. Hasil dan Pembahasan

Penelitian ini berhasil mengumpulkan 1754 data komentar dari video yang membahas tentang AI di YouTube. Setelah proses *pre-processing*, komentar-komentar ini diklasifikasikan menjadi dua kategori sentimen: positif (958) dan negatif (796).



Gambar 2. Proses Labeling Sentimen

Pada gambar 6 dengan menggunakan metode validasi 10-Fold Cross Validation, model ini mampu mencapai akurasi rata-rata 70%, yang menunjukkan bahwa algoritma Naive Bayes dapat bekerja cukup baik dalam analisis sentimen terhadap desainer grafis. Akurasi tertinggi yang dicapai adalah 76% pada fold kedua, dan akurasi terendah adalah 65% pada fold keenam.

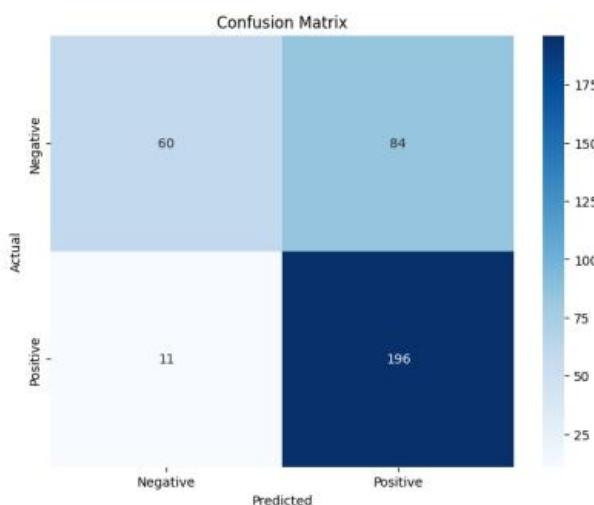


Gambar 3. Hasil dari 10 Fold Cross Validation

Pada table 1 Evaluasi menggunakan confusion matrix akan mendapatkan data seperti akurasi, presisi, recall, dan skor F1, dan memberikan gambaran menyeluruh tentang seberapa baik model Anda dalam melakukan prediksi terhadap setiap kelas target, Berikut tanel classification report.

Tabel 1
Hasil Classification Report

| | Percision | Recall | F1-score |
|--------------|-----------|--------|----------|
| Negatif | 0.85 | 0.42 | 0.56 |
| Positif | 0.70 | 0.95 | 0.80 |
| Accuracy | | | 0.73 |
| Macro avg | 0.77 | 0.68 | 0.68 |
| Weighted avg | 0.76 | 0.73 | 0.70 |



Gambar 4. Hasil Confusion Matrix

Hasil dari algoritma naïve bayes diatas menyatakan bahwa akurasi dari metode evaluasi di atas adalah 73%. Pada kelas precision dinyatakan juga bahwa prediction positif sebesar 70%, dan prediction negatif adalah 85%.

D. Kesimpulan

Setelah analisis sentimen yang dilakukan menggunakan algoritma Naive Bayes, beberapa poin penting dapat disimpulkan dari evaluasi kinerja model dengan confusion matrix dan validasi dengan cross-validation sepuluh kali. 73% adalah akurasi keseluruhan model yang diperoleh dari evaluasi confusion matrix, yang menunjukkan bahwa model dapat mengklasifikasikan sentimen dengan benar pada 73% dari data uji. Dengan menggunakan cross-validation sepuluh fold, validasi menunjukkan variasi dalam kinerja model di setiap fold. Fold terbaik adalah fold 2, yang memiliki akurasi sebesar 76%, fold terburuk adalah fold 6, yang memiliki akurasi sebesar 65%, dan rata-rata akurasi dari sepuluh fold adalah 70%. Hasil ini

menunjukkan bahwa para desainer grafis harus tetap mengikuti perkembangan zaman meskipun AI berkembang dengan sangat cepat dan berdampak besar pada dunia desain grafis.

Daftar Pustaka

- Kusuma Wardani, S., & Arum Sari, Y. (2021). *Analisis Sentimen menggunakan Metode Naïve Bayes Classifier terhadap Review Produk Perawatan Kulit Wajah menggunakan Seleksi Fitur N-gram dan Document Frequency Thresholding* (Vol. 5, Issue 12). <http://j-ptiik.ub.ac.id>
- Maulidah, N., Supriyadi, R., Utami, D. Y., Hasan, F. N., Fauzi, A., & Christian, A. (2021). Prediksi Penyakit Diabetes Melitus Menggunakan Metode Support Vector Machine dan Naive Bayes. *Indonesian Journal on Software Engineering (IJSE)*, 7(1), 63–68. <http://ejournal.bsi.ac.id/ejurnal/index.php/ijse63>
- Miranda, E., Gabriella, V., Wahyudi, S. A., & Chai, J. (2023). Text Classification untuk Menganalisis Sentimen Pendapat Masyarakat Indonesia terhadap Vaksinasi Covid-19. In *Kb. Jeruk, Kec. Kb. Jeruk*. <http://sistemasi.ftik.unisi.ac.id>
- Nugroho, A., & Religia, Y. (2021). Analisis Optimasi Algoritma Klasifikasi Naive Bayes menggunakan Genetic Algorithm dan Bagging. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(3), 504–510. <https://doi.org/10.29207/resti.v5i3.3067>
- Nur Fadilla, A., & Munadiyah Ramadhani, P. (2023). *Problematika Penggunaan AI (Artificial Intelligence) di Bidang Ilustrasi : AI VS Artist*. 4(1). <https://doi.org/10.33153/citrawira.v4i1.47>
- Sobron, M., Lubis Bidang, Y., Manufaktur, T., Teknik, P., Jurusan, M., Industri, T., & Kunci, K.-K. (2022). *IMPLEMENTASI ARTIFICIAL INTELLIGENCE PADA SYSTEM MANUFAKTUR TERPADU*.
- Utomo, A., Andadinata, M. A., & Widhiandono, D. (2023). ANALISIS KUALITAS KONTEN YOUTUBE BERDASARKAN KOLOM KOMENTAR DARI CHANNEL YOUTUBE BAIM PAULA. In *RELASI: Jurnal Penelitian Komunikasi* (Vol. 03, Issue 03).