

ANALISIS SENTIMEN DI TWITTER TERHADAP KENAIKAN UANG KULIAH TUNGGAL MENGGUNAKAN *ARTIFICIAL NEURAL NETWORK* (ANN)

Achmad Chikham Nouriel Rosyadi^{1*}, Katarina Rosa Leonida², Muhammad
Athoillah³, Hani Brilianti Rohmanto⁴

^{1,2,3,4} Program Studi Statistika, Universitas PGRI Adi Buana Surabaya, Indonesia

*Email: chikam015@gmail.com

Abstrak

Artificial Neural Network (ANN), khususnya metode Single Layer Perceptron (SLP), telah menunjukkan efektivitas dalam klasifikasi teks cepat dan akurat. Penelitian ini mengaplikasikan SLP untuk menganalisis sentimen masyarakat terhadap kenaikan UKT di Twitter. Hasil menunjukkan performa model yang baik dengan rata-rata akurasi, presisi, recall, dan F1-score sebesar 84%. Ini mengindikasikan kemampuan SLP yang cukup baik dalam mengklasifikasikan sentimen terhadap kebijakan Kemendikbud terkait kenaikan UKT. Penelitian menyimpulkan bahwa peningkatan jumlah data dapat lebih mengoptimalkan kinerja model dalam proses training dan meningkatkan akurasi prediksi pada data testing.

Kata kunci: *Single Layer Perceptron; Confusion Matrix; Kemendikbud*

Abstract

Artificial Neural Network (ANN), particularly the Single Layer Perceptron (SLP) method, has shown effectiveness in fast and accurate text classification. This research applies SLP to analyze public sentiment towards tuition fee increase on Twitter. The results show good model performance with an average accuracy, precision, recall, and F1-score of 84%. This indicates the ability of SLP which is quite good in classifying sentiment towards the Ministry of Education and Culture's policy regarding the UKT increase. The study concluded that increasing the amount of data can further optimize model performance in the training process and improve prediction accuracy on testing data.

Keywords: *Single Layer Perceptron; Confusion Matrix; Ministry of Education and Culture*

Copyright © (2024) Seminar Hasil Riset dan Pengabdian ke 6

PENDAHULUAN

Kenaikan Uang Kuliah Tunggal (UKT) di perguruan tinggi Indonesia telah menjadi isu kontroversial, menimbulkan kekhawatiran tentang akses pendidikan bagi mahasiswa kurang mampu. Dampaknya dapat berupa penundaan pendidikan hingga putus kuliah. Meskipun pemerintah telah merespons dengan Permendikbud Nomor 55 Tahun 2013 dan program bantuan, upaya ini masih dianggap belum memadai. Mahasiswa telah melakukan berbagai bentuk

protes, dan Twitter menjadi platform utama untuk menyuarakan opini, ditandai dengan viralnya hashtag seperti #TolakUKTMahal. Diduga terdapat Kesenjangan informasi antara pembuat kebijakan dan mahasiswa mengenai penerapan kenaikan UKT. Analisis sentimen dapat menjembatani kesenjangan ini dengan memberikan gambaran objektif tentang opini publik. melalui proses otomatis untuk memahami, mengekstrak, dan mengolah data tekstual guna mengungkap sentimen yang terkandung di dalamnya.

Analisis sentimen ini menggunakan *Artificial Neural Network* (ANN) dengan *metode Single Layer Perceptron* (SLP). SLP dipilih karena struktur sederhananya, kecepatan proses, dan efektivitasnya dalam klasifikasi biner (positif/negatif). Beberapa penelitian terdahulu telah menunjukkan keberhasilan SLP dalam berbagai konteks. Misalnya, penelitian (Desiani et al., 2023) menunjukkan bahwa SLP memiliki nilai presisi dan *recall* yang lebih baik dibandingkan algoritma *Adaboost* dalam klasifikasi penyakit Hepatitis-C. Penelitian (Aulia & Hermawan, 2023) juga mendemonstrasikan tingkat akurasi yang baik (81%) dalam analisis sentimen ulasan produk Tokopedia menggunakan algoritma *Perceptron*.

Tujuan penelitian ini adalah memahami opini masyarakat dalam konteks sosial-ekonomi kenaikan UKT pada platform twitter serta dilakukan prediksi sentimen menggunakan *Artificial Neural Network* (ANN) dengan *metode Single Layer Perceptron* (SLP). Diharapkan penelitian ini memberikan prediksi dampak sosial, perencanaan solusi yang lebih tepat sasaran, serta meningkatkan transparansi dan partisipasi publik dalam diskusi kebijakan pendidikan.

METODE

Metode penelitian yang digunakan untuk menganalisis sentimen di Twitter terhadap kenaikan UKT menggunakan algoritma *Single Layer Perceptron* (SLP), meliputi pengumpulan data, pembersihan, pelabelan data guna dilakukan sentimen analisis. Sentimen analisis dimana konsep untuk memprediksi kelas dari tweets yang label belum diketahui (Fitriana et al., 2021) dan prediksi sentimen tweets masyarakat.

Pengumpulan Data

Data dikumpulkan dari tweet dengan hashtag #kenaikanUKT sebanyak 400 tweet pada 22 Juni 2024, dengan menggunakan tools library Tweet-Harvest pada Python. Selanjutnya, data dilabeli secara manual menjadi kalimat positif dan negatif.

Preprocessing Data

Data Tweet yang sudah diambil masih berupa data mentah yang belum siap diolah, oleh karena itu dilakukan tahap *preprocessing* untuk mendapatkan data yang siap untuk. Tahapan *preprocessing* yang dilakukan sebagai berikut:

1. *Case folding*, tweet yang mengandung huruf kapital diubah menjadi huruf kecil.

2. *Cleansing*, menghapus tweet berisi tab, *new line*, *back slice*, *emoticon*, non ASCII, *mention*, *hashtag*, link, URL, angka, tanda baca dan whitespace.
3. *Tokenizing*, tweet yang berupa kalimat akan dipotong menjadi beberapa susunan kata.
4. *Normalizing*, tweet yang berisi kata gaul akan diubah menjadi kata baku.
5. *Stopword Removal*, menghapus tweet yang berisi kata tidak memiliki makna.
6. *Stemming*, tweet yang berisi kata imbuhan akan diubah menjadi kata dasar.

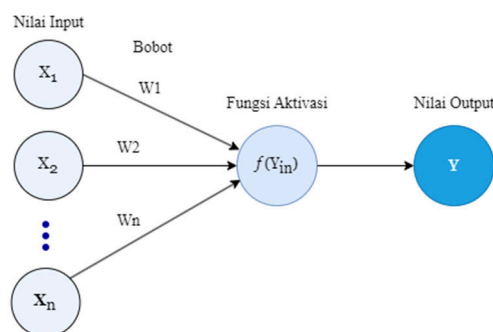
Term Frequency – Inverse Document Frequency (TF-IDF)

TF-IDF menghitung bobot dengan dua konsep, frekuensi kemunculan kata dalam dokumen dan kebalikan dari frekuensi dokumen yang mengandung kata tersebut. Frekuensi kemunculan kata menunjukkan seberapa penting kata tersebut dalam dokumen (Lestandy et al., 2021).

Single Layer Perceptron (SLP)

Single Layer Perceptron (SLP) adalah arsitektur *Artificial Neural Network* (ANN), terdiri dari dua komponen lapisan input dan output, tanpa lapisan tersembunyi. Umumnya diaplikasikan pada klasifikasi yang tidak terlalu kompleks, dengan fokus utama pada klasifikasi biner (Putri et al., 2023). Arsitektur SLP pada **Gambar 1**, dimana $X = [x_1, x_2, \dots, x_n]$ adalah input dari vektor angka, $W = [w_1, w_2, \dots, w_n]$ adalah vektor pembobot dan nilai *threshod* dari w_0 sebagai bias dan konstanta dari $x_0 = 1$. Kemudian fungsi penjumlahan atas perkalian bobot, W , dan nilai input, X , akan dilewatkan melalui fungsi aktivasi untuk menghasilkan output, $f(x) = y$. Fungsi ini digunakan untuk menghitung output y menggunakan fungsi aktivasi sigmoid dapat disajikan sebagai (Tela, 2020).

$$f(x) = y = \alpha \left(\sum_{i=0}^n x_i w_i \right) \quad (1)$$



Gambar 1. Arsitektur *Single Layer Perceptron*

Confusion Matrix

Confusion Matrix merupakan suatu representasi tabular yang dirancang untuk menggambarkan performa suatu algoritma klasifikasi secara visual, memungkinkan analisis

komprehensif terhadap akurasi prediksi model, menunjukkan jumlah klasifikasi yang benar dan salah untuk setiap kategori (Gunawan et al., 2021).

Tabel 1. Confusion Matrix

	Prediksi		
Realita	True Positive (TP)	False Negative (FN)	$\text{Recall} = \left(\frac{TP}{TP + FN}\right) \times 100\%$
	False Positive (FP)	True Negative (TN)	$\text{F1-Score} = 2 \times \left(\frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}}\right) \times 100\%$
	$\text{Presisi} = \left(\frac{TP}{TP + FP}\right) \times 100\%$	$\text{Akurasi} = \left(\frac{TP + TN}{(TP + TN + FP + FN)}\right) \times 100\%$	

HASIL DAN PEMBAHASAN

Pengumpulan Data

Pengumpulan data dilakukan melalui proses *crawling* dari platform Twitter menggunakan Google Colab, didapatkan sebanyak 400 tweets dengan *keyword* “kenaikan UKT”. Selanjutnya akan dilakukan proses pelabelan dengan mengisi label pada teks berisi Negatif (0) dan Positif (1) (Nursiyono & Chotimah, 2021).

Tabel 2. Pengumpulan Data Dan Pelabelan Data

Tweets	Label
Kenaikan UKT Kuliah Tahun 2024 https://t.co/2D0bLb3w10	1

Preprocessing Data

Proses data *preprocessing* dilakukan menggunakan program dengan bahasa pemrograman Python, library Sastrawi. Tahapan ini melalui *case folding*, *cleansing*, *tokenizing*, *normalizing*, *stopword removal* dan *stemming*, pada data hasil crawl sebanyak 400 tweet. Tahap pertama dilakukan yakni *case folding* untuk mengubah semua huruf kapital menjadi huruf kecil.

Tabel 3. Hasil Case Folding

Sebelum Case Folding	Sesudah Case Folding
Kenaikan UKT Kuliah Tahun 2024 https://t.co/2D0bLb3w10	kenaikan ukt kuliah tahun 2024 https://t.co/2d0bLb3wlo

Berdasarkan **Tabel 3** baris kedua terdapat hufu kapital yang sudah menjadi huruf kecil pada hasil *case folding*. Tahap kedua proses adalah *cleansing*.

Tabel 4. Hasil *Cleansing*

Sebelum Cleansing	Sesudah Cleansing
kenaikan ukt kuliah tahun 2024 https://t.co/2d0bLb3wlo	kenaikan ukt kuliah tahun

Berdasarkan **Tabel 4** hasil *cleansing* menunjukkan tanda baca yang ada sudah dihapus.

Tabel 5. Hasil *Tokenizing*

Sebelum Tokenizing	Sesudah Tokenizing
kenaikan ukt kuliah tahun	kenaikan, ukt, kuliah, tahun

Berdasarkan **Tabel 5** hasil *tokenizing* menunjukkan tweets sudah dipotong menjadi per kata.

Tabel 6. Hasil *Normalizing*

Sebelum Normalizing	Sesudah Normalizing
kenaikan, ukt, kuliah, tahun	kenaikan, ukt, kuliah, tahun

Berdasarkan **Tabel 6** hasil *normalizing* menunjukkan tidak ada tweets yang terdeteksi sebagai kata gaul.

Tabel 7. Hasil *Stopword Removal*

Sebelum Stopword Removal	Sesudah Stopword Removal
kenaikan, ukt, kuliah, tahun	kenaikan, ukt, kuliah, tahun

Berdasarkan **Tabel 7** hasil *Stopword Removal* sudah menghapus kata yang dianggap tidak memiliki makna.

Tabel 8. Hasil *Stemming*

Sebelum Stemming	Sesudah Stemming
kenaikan, ukt, kuliah, tahun	naik, ukt, kuliah, tahun

Berdasarkan **Tabel 8** hasil *Stemming*, tweet yang berisi kata imbuhan sudah diubah menjadi kata dasar.

Pembobotan TF-IDF

Pembobotan TF-IDF dilakukan secara terpisah sebelum melibatkan model SLP. Prosesnya melibatkan *tokenizing*, perhitungan TF, ID, kemudian produk TF-IDF menghasilkan matriks TF-IDF yang akan digunakan sebagai input untuk model SLP. Lalu model akan belajar dari matriks tersebut untuk dilakukan prediksi berdasarkan label yang sudah diberikan sebelumnya.

Hasil Algoritma Single Layer Perceptron (SLP)

Penelitian ini menggunakan *Single Layer Perceptron* (SLP) dengan membagi data menjadi 320 tweet untuk training dan 80 tweet untuk testing (rasio 80:20). Untuk mengurangi bias dan meningkatkan keandalan hasil, dilakukan 10 kali percobaan dengan kombinasi data training yang berbeda, dengan tetap menggunakan rasio 80:20. Kinerja model dinilai berdasarkan 4 nilai yakni akurasi, di mana menggambarkan persentase dari total prediksi yang benar oleh model. Lalu

presisi, yang mengukur proporsi prediksi positif yang benar-benar positif. *recall*, mengukur proporsi aktual positif yang diidentifikasi dengan benar oleh model. dan *F1-score*, merupakan rata-rata harmonis dari presisi dan recall. Kemudian dirata-rata dari 10 percobaan tersebut.

Tabel 9. Evaluasi Model

Percobaan	Akurasi	Presisi	Recall	F1-Score
1	86%	87%	85%	86%
2	83%	81%	83%	82%
3	83%	84%	83%	84%
4	87%	87%	87%	87%
5	86%	85%	86%	85%
6	81%	82%	81%	83%
7	85%	83%	85%	83%
8	81%	79%	80%	79%
9	82%	81%	82%	82%
10	87%	86%	87%	86%
Rata-rata	84%	84%	84%	84%

Berdasarkan **Tabel 9** hasil evaluasi menunjukkan secara keseluruhan, nilai rata-rata 84% untuk akurasi, presisi, *recall*, dan *F1-score* menunjukkan bahwa model SLP yang digunakan untuk prediksi sentimen analisis memiliki kinerja yang baik. Model ini cukup efektif dalam mengklasifikasikan tweet dan prediksi yang benar. Artinya, model SLP dapat diandalkan dalam analisis sentimen masyarakat terkait kenaikan UKT di platform Twitter. Kinerja model juga dapat dilihat dari hasil evaluasi model dari 10 percobaan yang dilakukan. Evaluasi model dengan persentase tertinggi didapatkan pada percobaan ke-4, model mencapai akurasi 87% menunjukkan bahwa model dapat mengklasifikasikan tweet dengan benar secara keseluruhan. Nilai presisi yang tinggi menunjukkan bahwa model memprediksi tweet sebagai positif, kemungkinan besar prediksi tersebut benar dengan presisi mencapai 87%. Nilai *recall* model dapat menangkap sebagian besar tweet positif yang ada mencapai 87%. *F1-Score* sebesar 87%, menunjukkan bahwa model tidak hanya baik dalam menangkap tweet positif, tetapi juga menghindari kesalahan prediksi positif palsu mencapai. Sementara evaluasi model dengan nilai terendah pada percobaan ke- 8, dengan nilai akurasi 81%, presisi 79%, *recall* 80% dan *F1-score* 79%. Variasi dalam hasil percobaan menunjukkan bahwa model konsisten namun masih memiliki ruang untuk perbaikan, terutama dalam meningkatkan *recall* untuk menangkap lebih banyak tweet positif tanpa mengorbankan presisi, hal ini diakibatkan karen ketidakseimbangan data tweets dengan label negatif yang lebih banyak daripada positif.

Kesimpulan

Keseluruhan hasil uji coba ini menunjukkan hasil analisis bahwa model SLP dalam melakukan klasifikasi sentimen analisis masyarakat pada kebijakan kemendikbud dalam kenaikan UKT memiliki kinerja yang cukup baik, ditunjukkan dengan nilai rata-rata akurasi, presisi, *recall*, dan *f1-score* sebesar 84%. Perlu diperhatikan bahwa, model SLP ini dapat dibangun

dengan jumlah data yang cukup banyak agar dalam proses *training* dapat berjalan dengan optimal dan data *testing* dapat menghasilkan prediksi yang akurat.

DAFTAR PUSTAKA

- Aulia, M., & Hermawan, A. (2023). Analisis Perbandingan Algoritma SVM, Naïve Bayes, dan Perceptron untuk Analisis Sentimen Ulasan Produk Tokopedia. *Jurnal Media Informatika Budidarma*, 7(4), 1850–1859. <https://doi.org/10.30865/mib.v7i4.6839>
- Desiani, A., Maiyanti, I., Suprihatin, B., Kurniawan, R., Adzra, D., Nabila, A., Matematika, J., Matematika, F., & Alam, I. P. (2023). IMPLEMENTASI ALGORITMA ADAPTIVE BOOSTING (ADABOOST) DAN SINGLE LAYER PERCEPTRON (SLP) PADA KLASIFIKASI PENYAKIT HEPATITIS-C Implementation of Adaptive Boosting (Adaboost) and Single Layer Perceptron (SLP) Algorithm in Hepatitis-C Disease Classification. *ScientiCO : Computer Science and Informatics Journal*, 6(2).
- Fitriannah, D., Dwiasnati, S., H, H. H., & Baihaqi, K. A. (2021). Penerapan Metode Machine Learning untuk Prediksi Nasabah Potensial menggunakan Algoritma Klasifikasi Naïve Bayes. *Faktor Exacta*, 14(2), 92. <https://doi.org/10.30998/faktorexacta.v14i2.9297>
- Gunawan, A. L. S. A.-Z., Jondri, & Lhaksamana, K. M. (2021). Analisis Sentimen pada Media Sosial Twitter terhadap Penanganan Bencana Banjir di Jawa Barat dengan Metode Jaringan Saraf Tiruan Sentiment. *E-Proceeding of Engineering*, 8(2), 2965. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/viewFile/14695/14472>
- Lestandy, M., Abdurrahim, A., & Syafa'ah, L. (2021). Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(10), 802–808.
- Moldwin, T., & Segev, I. (2020). Perceptron Learning and Classification in a Modeled Cortical Pyramidal Cell. *Frontiers in Computational Neuroscience*, 14(April), 1–13. <https://doi.org/10.3389/fncom.2020.00033>
- Nursiyono, J. A., & Chotimah, C. (2021). Analisis Sentimen Netizen Twitter terhadap Pemberitaan PPN Sembako dan Jasa Pendidikan dengan Pendekatan Social Network Analysis dan Naive Bayes Classifier. *J Statistika: Jurnal Ilmiah Teori Dan Aplikasi Statistika*, 14(1), 52–58. <https://doi.org/10.36456/jstat.vol14.no1.a3868>
- Putri, R. K., Athoillah, M., Haqiqiyah, A., & Lestari, F. W. A. (2023). Deteksi Penggunaan Masker Wajah Dengan Algoritma Deep Learning. *Snhrp*, 2023. <https://snhrp.unipasby.ac.id/prosiding/index.php/snhrp/article/view/751>
- Sinaga, L., Irawan, E., Saputra, W., Sudahri Damanik, I., Syahputra Saragih, I., Tunas Bangsa Pematangsiantar, S., & Jendral Sudirman Blok No, J. A. (2020). Penerapan Jaringan Saraf Tiruan Metode Backpropogation Dalam Memprediksi Distribusi Air Pada PDAM Tirtauli Kota Pematangsiantar. *Prosiding Seminar Nasional Riset Dan Information Science (SENARIS)*, 2, 161–168.
- Tela, A. F. (2020). *Sentiment Analysis for Low-Resource Language : The Case of Tigrinya*. May.