

KLASIFIKASI SENTIMEN TWITTER MENGGUNAKAN LSTM

Firman Yuspriyadi

Program Studi Magister Ilmu Komputer, Universitas Nusa Mandiri

Jl. Kramat Raya No. 25 10450 Jakarta Pusat, Jakarta, Indonesia

Email: 14210223@nusamandiri.ac.id

ABSTRACT

One of the fastest growing social media users is Twitter, the number of Twitter users is said to continue to increase by 300,000 users every day. Twitter users send twitter posts (tweets) regarding facts and opinions on government products or services they use or express their political, ideological views and interests. It is no exception to sending opinion tweets regarding influential leaders or public figures in this country. With 55 million tweets every day, Twitter has a high update rate and is a very efficient data warehouse for research in the political and social fields, so Twitter is a good place to do opinion mining or sentiment analysis in classifying tweets. This study aims to classify tweets based on positive and negative sentiments using the LSTM method. The dataset used comes from <https://github.com/riochr17/Analisis-Sentimen-ID>. Based on the test results show that the LSTM method has an accuracy of 0.539, F1 Score 0.7

Keywords- LSTM, Machine Learning, Twitter, Word2Vec

I. PENDAHULUAN

Mikroblog seperti Twitter (www.twitter.com) dan Facebook (www.facebook.com) sekarang menjadi perangkat komunikasi yang sangat populer kalangan pengguna internet. Salah satu media sosial yang berkembang sangat pesat penggunaannya adalah Twitter, jumlah pengguna twitter disebutkan terus meningkat 300.000 user setiap harinya. Statistik tersebut menyebutkan bahwa pada bulan Februari 2022, twitter memiliki 217 juta akun[1]. Setiap harinya para pengguna twitter mengirimkan twitter post (tweet) mengenai fakta dan opini produk atau layanan pemerintah yang mereka gunakan atau mengekspresikan pandangan politik, ideologis dan minat mereka. Tidak terkecuali juga mengirimkan tweet opini terkait pemimpin atau tokoh publik yang berpengaruh di negara ini. menyediakan sumber-sumber opini yang besar jumlahnya bagi kebutuhan individu maupun organisasi. Melalui media sosial orang dapat mengekspresikan apa saja, termasuk pendapatnya akan sesuatu hal tanpa adanya keterpaksaan.

Menjadikan twitter memiliki tingkat update yang tinggi. Hal inilah yang mengakibatkan tingginya ketersediaan data di twitter, sehingga twitter merupakan tempat yang baik untuk melakukan opinion mining atau analisis sentimen. Gudang data yang ada di twitter inilah yang sangat efisien untuk penelitian dibidang politik, pemasaran dan sosial.

Secara umum terdapat dua tipe informasi tekstual di web yaitu fakta dan opini. Fakta adalah pernyataan objektif mengenai entitas dan kejadian di dunia sedangkan opini adalah pernyataan subjektif yang merefleksikan sentimen atau persepsi orang mengenai entitas ataupun kejadian di dunia. Opini akan menjadi penting ketika calon Gubernur DKI Jakarta Tahun 2017 ingin membuat sebuah keputusan dengan terlebih dahulu mendengar opini dari pihak lain. Sekarang ketika calon gubernur DKI Jakarta tahun 2017 ingin memperoleh opini publik mengenai produk, citra dan

layanannya, maka mereka tidak perlu melakukan survey konvensional dan focus group yang lama dan mahal biayanya.

Di beberapa Negara, Twitter telah dimanfaatkan untuk menjangkau pendapat masyarakat terhadaptokoh publik dan juga prediksi calon legislatif yang akan terpilih seperti Singapura, Jerman dan Amerika. Hal itu dikarenakan twitter merupakan salah satu media jejaring sosial dengan pengguna terbanyak diantara beberapa situs jejaring sosial yang ada. Hal ini juga menimbulkan motivasi khusus untuk mengembangkan suatu penelitian. Penelitian ini membahas tentang bagaimana melakukan klasifikasi dengan menganalisis sentimen pada opini tweet berbahasa indonesia pada calon Gubernur DKI Jakarta tahun 2017 dengan mengekstraksi fitur menggunakan Word2Vec. Algoritma untuk melakukan klasifikasi menggunakan algoritma Long-Short Term Memory (LSTM).

II. TINJAUAN PUSTAKA

Berbeda dari rule base, machine learning adalah metode pembelajaran mesin (komputer) yang menggunakan algoritma tertentu untuk melakukan tugas tanpa perlu menerapkan rule secara manual. Sejumlah data yang telah diberikan label diberikan sebagai input untuk computer menemukan pola (model) yang kemudian akan digunakan terhadap data baru. Berdasarkan perkembangannya, machine learning juga dapat dikelompokkan menjadi dua yaitu conventional machine learning dan deep learning [2]. Conventional Machine Learning memiliki struktur yang simple seperti linear regression atau decision tree sedangkan deep learning memiliki struktur yang kompleks seperti jaringan saraf tiruan .

Metode machine learning konvensional yang cukup efektif antara lain adalah support vector machine serta naïve bayes. Metode Naïve bayes yang digunakan pada [3] untuk tugas klasifikasi sentimen vaksin Covid-19 pada twitter. Tuning parameter, kombinasi teks preprocessing serta penyeimbangan jumlah data latih

pada dataset yang tidak seimbang. Menghasilkan akurasi tertinggi F1-score 57.15% dengan akurasi 61%. Metode support vector machine dengan input Bag of Word berdasarkan TF-IDF pada [4] dengan dataset yang sama, menghasilkan F1-score 56.81% dan akurasi 65% setelah tuning parameter, dan kombinasi setup eksperimen yang paling optimal.

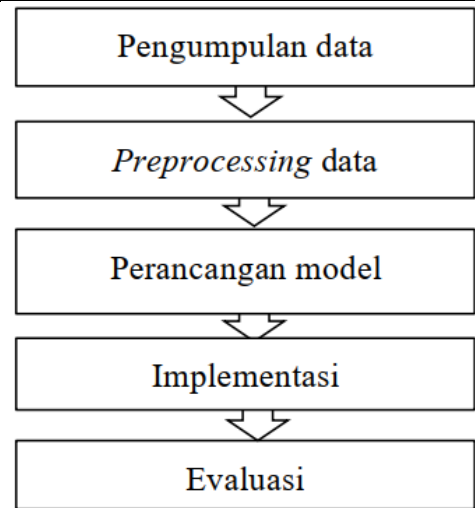
Metode machine learning konvensional memiliki masalah yaitu sulitnya untuk menentukan feature extraction yang harus dimasukkan pada model. Jika feature hilang atau belum komplit maka model akan memberi keluaran yang tidak sempurna. Seiring dengan perkembangan pemodelan dengan menggunakan Jaringan Syaraf Tiruan (JST), pada tahun 1997 muncul model baru yang disebut Long Short Term Memory (LSTM). LSTM sangat baik dalam menangani klasifikasi teks karna model LSTM dirancang untuk bekerja pada untuk menangani masalah ketergantungan jangka panjang dan merupakan unit spesial dari Recurrent Neural Network (RNN) [5].

Long short term memory (LSTM) yang digunakan untuk klasifikasi sentimen pada media sosial multiclass [6]. Hasil akurasi tertinggi diperoleh dari uji coba sebanyak 5 kali, adalah 91,9%. Word2vec sebagai input diterapkan pada sentimen ulasan hotel [7], dengan membandingkan model Continous Bag of Word (CBOW) serta word2vec Skip-gram. Hasil akurasi tertinggi 85.96% diperoleh pada model skip-gram.

Paper ini membahas metode LSTM dengan input word embeddings word2vec yang bertujuan untuk mengklasifikasi sentimen masyarakat pada pemilihan gubernur tahun 2017 di Twitter dan model dapat memprediksi kata kata baru. Eksperimen dilakukan terhadap dataset yang digunakan dan teks preprocessing yang sama untuk mengevaluasi apakah metode Long Short Term Memory dengan menggunakan Word2vec sebagai word embedding dapat memiliki performa yang lebih baik dari kedua metode machine learning konvensional itu.

III. METODOLOGI PENELITIAN

Dalam penelitian ini, ada beberapa tahap yang harus dilewati dengan baik. Tahapan penelitian yang dilakukan meliputi pengumpulan data berupa data saham dan sentimen publik, *preprocessing* data, perancangan, implementasi dan evaluasi. Diagram Alir penelitian dapat dilihat pada Gambar 1.



Gambar 1. Persentase Penyebab Kehilangan Data

a. Pengumpulan Data

Mengumpulkan data, yaitudata twitter berisipini masyarakat yang mengandung unsur nama calon Gubernur DKI jakarta tahun 2017 dari situs <https://github.com/riochr17/Analisis-Sentimen-ID>.

b. Pelabelan

Data tweet diberikan label berdasarkan isi tweetnya, apakah mengandung opini positif, negatif atau netral. Pelabelan opini ditunjukkan pada Gambar 2

	tweet	category
0	rt @napqilla: no 1, 3 ambisinya menguasai raky...	1
1	rt @pandji: nah gue pikir sentimen petahana ok...	1
2	rt @pandji: urutan pertama best moment #debat2...	1
3	rt @pandji: ini artikel yg menjelaskan ternyata...	1
4	rt @mrtampi: agus makin santai.\nahok makin sa...	0

Gambar 2. Tweet

c. Preprocessing

Tahap text preprocessing bertujuan untuk menyiapkan data sebelum proses klasifikasi [14]. Pada penelitian ini dilakukan beberapa tahapan preprocessing, yaitu hyperlink removal (menghapus hyperlink) dan pengubahan setiap mention (misalnya "@budi0232", "@supermania", dan lainnya) menjadi token "@USER". Selanjutnya penghapusan angka dan emoticon, lalu case folding (mengubah isi teks menjadi huruf kecil semuanya), dan punctuation removal (menghapus tanda baca). Proses stemming (pengembalian kata ke bentuk akar kata) dan stopword removal (menghapus kata-kata yang tidak penting) dilakukan dengan menggunakan menggunakan library sastrawi. Komposisi preprocessing yang digunakan ini mengacu pada pengalaman hasil optimal yang dicapai.

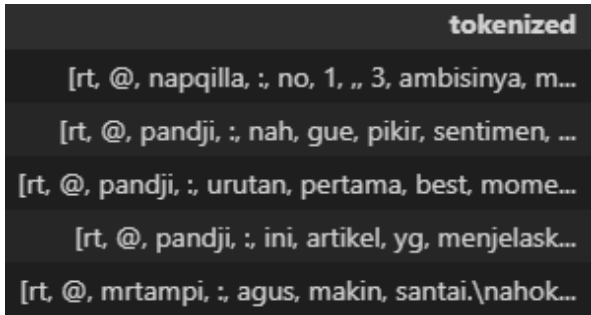
d. Training Word2Vec

Word2vec diperkenalkan oleh Mikolov, Corado & Chen [8]. Word2vec membantu sistem untuk mengenal kata yang direpresentasikan dalam bentuk vektor (*word embeddings*). *Word embeddings* sangat bermanfaat dalam tugas-tugas NLP dalam memperkecil dimensi vektor kata dibandingkan menggunakan *model bag of words*, agar lebih mudah dan efisien saat komputasi. Penelitian menyelidiki penggunaan word2vec dengan model pretrained Bahasa Indonesia (dilatih terlebih dahulu oleh periset

lain). Model tersebut dibangkitkan dari hasil latihan word2vec pada arsip data dump wikipedia pada tahun 2020. Model ini menjadi input untuk vektor kata-kata penyusun tweet pada LSTM.

e. Tokenisasi

Proses tokenisasi yang dilakukan (Tabel 3) adalah memecah-mecah teks (tweet) menjadi token (kata). Setiap kata akan menjadi input pada setiap node di dalam LSTM, hasil tokenisasi ditunjukkan pada gambar 3



Gambar 3. Hasil tokenisasi

f. Pad Sequences

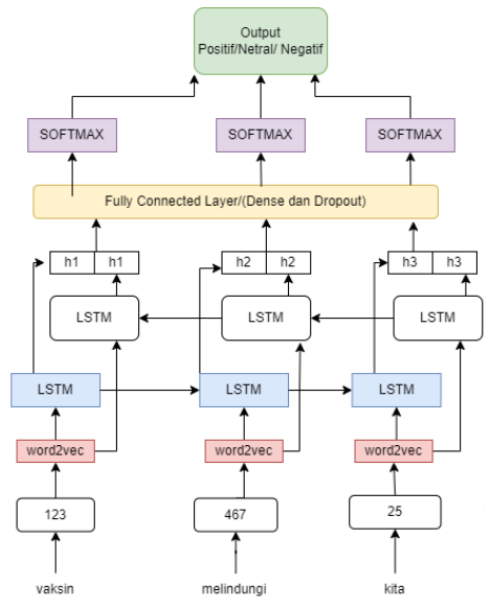
Dalam arsitektur deep learning, dimensi input harus konsisten untuk seluruh tweet di dalam dataset. Karena panjang tweet tidak sama, maka diperlukan suatu teknik untuk memenuhi jumlah node input pada LSTM. Pad Sequences adalah teknik yang biasa digunakan untuk kebutuhan tersebut. Suatu array dengan dimensi n dibentuk untuk menampung kata-kata pada tweet yang diwakili dengan nomor ID kata. Apabila jumlah kata lebih kecil dari n, maka node yang kosong akan diisi dengan nilai nol, sedangkan apabila panjang tweet melebihi n kata, maka sisanya akan dibuang [9]. Nilai nol dapat diisi pada bagian awal (padding awal), dan ID kata-kata tweet pada bagian akhir, atau sebaliknya. Penelitian ini menggunakan metode padding awal.

d. Proses Training LSTM

Metode yang digunakan adalah LSTM (Gambar 4). LSTM memanfaatkan konteks sebelumnya (Forward Layer) dan konteks setelahnya (Backward Layer) dengan memproses data dari dua arah dengan hidden layer terpisah. Proses training LSTM dimulai dari vektor input dari hasil sub bab 3.3 yang diubah ke dalam bentuk vektor word embeddings oleh blok vectorizer (fase model klasifikasi pada Gambar 3). Training dilakukan terhadap data latih, dengan validasi menggunakan dataset validasi. Proses latih dilakukan dengan epoch sebanyak 50 dan batch 64. Parameter tuning dilakukan untuk menemukan model klasifikasi yang paling optimal. Dalam penelitian ini, diambil metrik akurasi untuk menemukan model optimal dari klasifikasi menggunakan LSTM ini.

Proses LSTM (Gambar 4) dimulai dari memecahkan kalimat menjadi kata dan dijadikan sebagai indeks (tokenisasi). Kata yang telah token selanjutnya bentuk indeks tadi akan dijadikan bentuk vektor untuk menjadi inputan pada LSTM dengan word2vec. Selanjutnya, vektor word2vec masuk

kedalam proses Bidirectional LSTM yaitu forward dan backward layer untuk memproses data dari dua arah dengan hidden layer terpisah.



Gambar 4 Proses Training LSTM

IV. HASIL DAN PEMBAHASAN

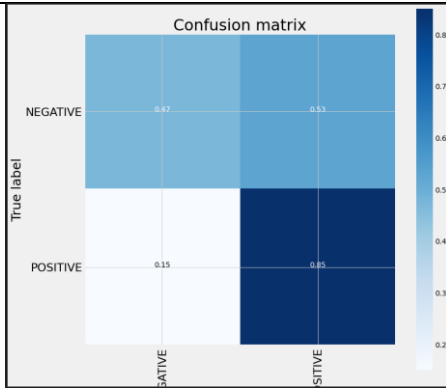
Tahap ini adalah tahap pencarian model optimal. Tahap ini memiliki tiga langkah yaitu : data balancing (penyeimbangan data), parameter tuning dan pengujian dengan data uji

4.1. Training Data

Proses pelatihan LSTM dengan data yang baru memberikan hasil F1-score yang meningkat signifikan, khususnya pada kelas negatif dan positif, seperti terlihat pada Gambar 5. Terlihat bahwa saat ini sistem telah dapat mengenali atau mendeteksi sentiment positif dan negatif. Perbandingan akurasi antara proses validasi untuk model awal dan model dengan balanced dataset, dapat dilihat pada gambar 5 berikut ini.

	precision	recall	f1-score	support
NEGATIVE	0.78	0.47	0.59	161
POSITIVE	0.59	0.85	0.69	141
accuracy			0.65	302
macro avg	0.68	0.66	0.64	302
weighted avg	0.69	0.65	0.64	302

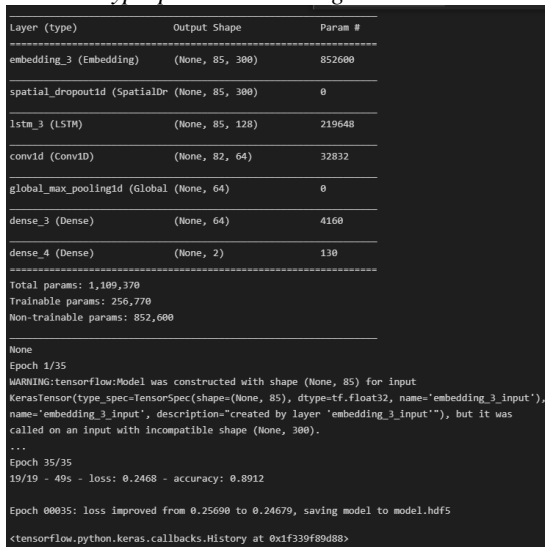
Gambar 5 Performa model dengandata latih awal



Gambar 6 Akurasi dalam *confusion matrix*

4.2 Parameter Tuning

Penyetelan parameter LSTM (parameter tuning) dilakukan untuk menemukan model yang paling optimal. Adapun parameter untuk dilakukan *hyperparameter tuning* adalah dengan menggunakan fungsi aktivasi yang dipilih adalah di antara fungsi RELU dan SOFTMAX dan *optimizer* Adam. Proses pelatihan menggunakan mekanisme early stopping untuk mendeteksi overfitting. Pada saat terjadi overfitting, sistem akan memanggil kembali (callback) model dengan hasil epoch sebelumnya yang nilai akurasi terhadap data validasinya lebih baik. Gambar 7 memperlihatkan hasil dari proses *training* setelah dilakukan *hyperparameter tuning*.



Gambar 7 Hasil *training hyperparameter tuning*

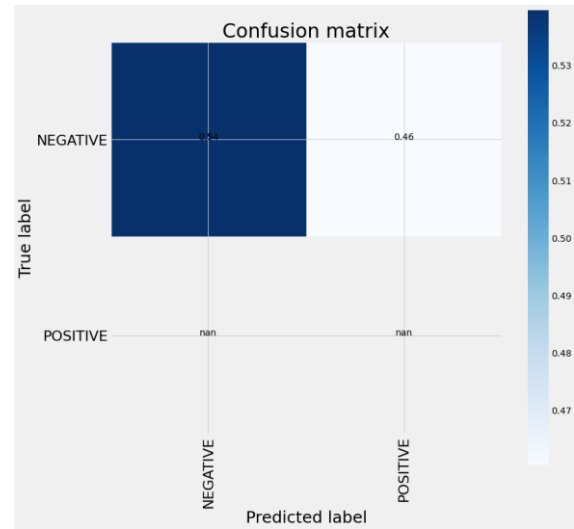
4.3 Pengujian Data Uji

Model final yang diperoleh, diterapkan pada data testing, yang belum pernah terlihat sebelumnya pada saat proses training klasifikasi menggunakan LSTM maupun training pembentukan language model menggunakan word2vec.

Dari hasil tersebut, metode LSTM yang diusulkan memiliki performa yang kompetitif bila ditinjau dari F1-score, namun dengan *score* akurasi 0,53, sebagaimana dapat dicermati pada gambar dibawah ini.

```
WARNING:tensorflow:Model was constructed with shape (None, 85) for input
KerasTensor(type_spec=TensorSpec(shape=(None, 85), dtype=tf.float32, name='embedding_3_input'),
name='embedding_3_input', description='created by layer 'embedding_3_input'), but it was
called on an input with incompatible shape (None, 300).
10/10 [-----] - 25 152ms/step
Test Accuracy 0.5397358993377483
F1 Score on Test Set: 0.7010752688172043
[[163 139]
 [ 0  0]]
```

Gambar 8 Hasil akurasi dan F1 Score



Gambar 9 Confusion matrix

Kurangnya jumlah data latih untuk kelas tweet yang mengandung sentiment, menyebabkan sistem kurang dapat mendeteksi kelas yang berisi sentiment positif dan negatif. Diperlukan lebih banyak kata-kata spesifik yang berisi sentiment untuk terlihat pada saat training, sehingga untuk ini, diperlukan lebih banyak data tweet berisi sentiment positif dan negatif, agar model hasil pelatihan lebih optimal lagi

V. KESIMPULAN DAN SARAN

Hasil implementasi serta pengujian yang telah dilaksanakan, bahwasannya metode Long Short Term Memory dapat diterapkan pada klasifikasi sentimen pada media sosial Twitter.. Pada kasus ini LSTM mempunyai akurasi 0.539, F1 Score 0.7. Nilai akurasi yang didapatkan tidak terlalu baik dikarenakan dataset yang sedikit dan tidak seimbang dan kasus yang dilakukan bukan klasifikasi biner, melainkan multilabel.

Saran untuk penelitian selanjutnya, adalah dengan menambah koleksi data tweet yang berisi sentiment positif dan negatif sehingga jumlah data yang banyak diharapkan dapat meningkatkan hasil klasifikasi menggunakan metode deep learning LSTM. Selain itu, masih terbuka peluang untuk menyelidiki model arsitektur LSTM yang lebih kompleks dan melakukan parameter tuning untuk model tersebut.

IV. REFERENSI

[1] Republika Online. Tersedia: <https://www.republika.co.id/berita/r73bky368/jumlah-pengguna-harian-twitter-217-juta-orang>

- [2] N. K. Chauhan and K. Singh, "A review on conventional machine learning vs deep learning," *Int. Conf. Comput. Power Commun. Technol. GUCON 2018*, pp. 347–352, 2019.
- [3] P. Yohana, "Analisis Sentimen Vaksin Covid19 Menggunakan Naive Bayes," Skripsi, 2022.
- [4] M. Rizky, "Vaksin Covid-19 Menggunakan Metode Support Vector Machine Pada Media Sosial Twitter Covid-19", Skripsi, 2021..
- [5] A. Arfan and L. ETP, "Perbandingan Algoritma Long Short-Term Memory dengan SVR Pada Prediksi Harga Saham di Indonesia," *Petir*, vol. 13, no. 1, pp. 33–43, 2020,.
- [6] Y. Astari and S. W. Rozaqi, "Analisis Sentimen Multi-Class pada Sosial Media menggunakan metode Long Short-Term Memory (LSTM)," vol. 4, no. 1, pp. 8–12, 2021..
- [7] P. F. Muhammad, R. Kusumaningrum, and A. Wibowo, "Sentiment Analysis Using Word2vec and Long Short-Term Memory (LSTM)
- [8] P. F. Muhammad, R. Kusumaningrum, and A. Wibowo, "Sentiment Analysis Using Word2vec and Long Short-Term Memory (LSTM) for Indonesian Hotel Reviews," *Procedia Comput. Sci.*, vol. 179, pp. 728–735, 2021
- [9] Y. Xie, L. Le, Y. Zhou, and V. V. Raghavan, *Deep Learning for Natural Language Processing*, vol. 38. 2018