

Klasifikasi Respons Terhadap Vaksinasi Covid-19 Berdasarkan *Tweets* Menggunakan *Attention-Based Long Short-Term Memory*

Diva Zannuba S, Bandung Arry, dan Muhammad Luthfi Shahab
Departemen Matematika, Institut Teknologi Sepuluh Nopember (ITS)
e-mail: bandung@matematika.its.ac.id; luthfishahab@matematika.its.ac.id.

Abstrak—Media sosial memudahkan masyarakat dalam mendapatkan informasi dan menuangkan pendapat, saran atau kritiknya dalam peristiwa tertentu. Vaksinasi virus COVID-19 di Indonesia yang sedang hangat diperbincangkan dan mendapatkan beragam respons dari masyarakat baik pro maupun kontra, dapat dimanfaatkan untuk melakukan analisis terhadap respons tersebut. Untuk mendukung analisis tersebut, dilakukan klasifikasi respons dari masyarakat Indonesia terhadap vaksinasi COVID-19 menjadi tiga kelas yaitu negatif, netral, dan positif. Untuk proses klasifikasi respons tersebut, diimplementasikan metode *Attentional-based Long Short Term Memory* atau A-LSTM. Disisi lain, penelitian ini juga mengimplementasikan *Bidirectional Encoder Representation Transformer* (BERT) sebagai metode pada proses tokenisasi untuk memperoleh representasi fitur dari data *Tweet* sehingga membantu proses pelatihan A-LSTM. Proses evaluasi dilakukan dengan menggunakan dataset *Tweets Bahasa Indonesia* dari media sosial *Twitter* dimulai dari diangkatnya isu vaksinasi COVID-19 di Indonesia. Hasil dari metode ini menunjukkan kinerja yang baik dengan nilai akurasi sebesar 82%.

Kata Kunci—Analisis Sentimen, *Attention-based Long Short Term Memory* (A-LSTM), Media Sosial, *Twitter*.

I. PENDAHULUAN

Di era digital, masyarakat sangat sulit untuk dipisahkan dengan media sosial. Manfaat media sosial tidak hanya memudahkan untuk mendapatkan informasi terkini melainkan juga dapat menampung pendapat, saran atau kritik dari *user* (pengguna). Begitu pula, *user* mampu mengekspresikan keadaan atau kondisi terkininya ataupun sekitar terhadap kejadian yang tengah diperbincangkan. Tercatat pada tahun 2020[1], sebanyak 198.96 juta jiwa penduduk Indonesia aktif menggunakan media sosial. Sedangkan, dalam artikel menyebutkan bahwa *Twitter* menjadi salah satu media sosial yang digemari masyarakat di Indonesia dengan jumlah *user* aktif yang tercatat mencapai 14.05 juta sampai Januari 2021. *Twitter* sendiri merupakan salah satu layanan media *on-line* dimana *user* diberikan akses untuk membagikan pesan, saran atau kritik dalam bentuk foto, video, tautan dan teks dengan batasan maksimal hingga 280 karakter yang sering dikenal dengan *Tweet*. Didalam *Twitter* juga terdapat beberapa fitur seperti *hashtag* dan *emoticon* yang digunakan untuk mengekspresikan respons dalam *Tweet* dari *user*.

Kasus pandemi COVID-19 yang ditemukan semenjak tahun 2019, terlebih vaksinasi COVID-19 di Indonesia menjadi salah satu kondisi yang ramai dibicarakan, terutama

sejak bulan Desember 2020 hingga awal tahun 2021. Pada saat itu, vaksin Sinovac telah didatangkan ke Indonesia dari negara Cina sekitar 3 juta dosis sebagai langkah penanggulangan pandemi COVID-19 yang akan didistribusikan ke masyarakat dimulai dari Presiden Republik Indonesia pada tanggal 13 Januari 2021. Distribusi vaksin ini menyebabkan pro dan kontra dari masyarakat Indonesia. Reaksi masyarakat yang beragam dapat berpengaruh terhadap kesediaan masyarakat untuk menerima vaksin, juga lancar tidaknya program vaksinasi COVID-19 di Indonesia. Untuk mengetahui hal ini, Kementerian Kesehatan melakukan survei tentang vaksinasi COVID-19 kepada masyarakat yang telah diisi hingga 115.000 responden dimana hasil respons dapat disimpulkan bahwa sebagian besar menyatakan bersedia menerima vaksin[2].

Twitter sebagai salah satu sumber informasi dapat digunakan sebagai cara lain untuk mengetahui respons masyarakat, karena masyarakat dapat mengutarakan opininya terkait vaksinasi COVID-19 di Indonesia dengan lebih bebas. *Twitter* dapat menjadi tempat untuk mengumpulkan berbagai respons masyarakat, yang kemudian dapat dilakukan klasifikasi terhadap respons yang telah terkumpul. Maka diperlukan penelitian untuk mengklasifikasikan respons sentimen ini sehingga dapat digunakan untuk menganalisa respons masyarakat Indonesia terhadap vaksinasi COVID-19. Respons yang dimaksudkan pada penelitian ini dapat diartikan kedalam sentimen seperti positif, negatif, dan netral. Ada beberapa metode yang biasa digunakan untuk melakukan penelitian sentimen, seperti *machine learning* dan *deep learning*. Penelitian serupa telah dilakukan Aryal[3] yang telah melakukan analisis sentimen terkait vaksinasi COVID-19 dari *tweet* berdasarkan lokasi dari negara Nepal, India, dan Singapura menggunakan metode *machine learning* Na'ive Bayes dan metode *deep learning Long Short-Term Memory* (LSTM). Hasil dari penelitian ini menandakan bahwa metode LSTM lebih unggul daripada metode Na'ive Bayes, sehingga metode LSTM akan diterapkan pada penelitian ini. Penelitian ini juga akan menggunakan *tweet* berbahasa Indonesia, yang menjadi salah satu pembeda dari penelitian diatas.

Beberapa penelitian dengan metode serupa dilakukan oleh Imran dkk [4] yang telah melakukan deteksi emosi dan analisis sentimen pada *Tweets* selama masa pandemi COVID-19 dengan menggunakan metode LSTM. Hasil dari penelitian ini sudah cukup bagus. Akan tetapi, banyak juga penelitian pada bidang *Natural Language Processing* (NLP) yang menambahkan *attention mechanism* pada LSTM, supaya dapat menangkap representasi fitur teks dengan lebih

baik. Seperti yang digunakan oleh Yuxio Chen dkk [5], yang telah melakukan analisis sentimen positif dan negatif pada *Twitter* menggunakan *bi-sense emoji embedding* dan *Attention-based LSTM Network* dengan nilai akurasi yang lebih tinggi daripada penelitian milik Imran. Namun penelitian ini lebih memberikan perhatian yang lebih besar ke emoji yang digunakan, sehingga pada penelitian ini hanya akan berfokus terhadap teks pada *tweet* dan tidak menggunakan metode *bi-sense emoji embedding* tersebut.

Berdasarkan paparan diatas, maka pada penelitian ini akan mengklasifikasikan dan menganalisa respons masyarakat Indonesia terkait vaksinasi COVID-19 dengan menerapkan Attention Mechanism ke dalam LSTM, yang dapat disebut dengan Attention-based LSTM atau A-LSTM. Sebagai tambahan, pada penelitian ini juga menggunakan metode Bidirectional Encoder Representation Transformer (BERT) sebagai salah satu proses dalam melakukan penelitian NLP.

II. PENELITIAN TERDAHULU

Beberapa penelitian terkait klasifikasi dari komentar di media sosial, diantaranya adalah penelitian yang dilakukan Aryal dan Bhattarai [6], yang melakukan analisis sentimen terhadap respons masyarakat terhadap vaksinasi COVID-19 di wilayah Nepal, India, dan Singapura. Penelitian ini menggunakan dua metode yaitu Naïve Bayes dan LSTM, dengan akurasi dari kedua metode yaitu 77,25% dan 84,13%. Junianto dan Rachman [6] melakukan deteksi emosi dari komentar pada media sosial dengan menggunakan *Naive Bayes Classifier*, yang mengklasifikasikan empat emosi senang, sedih, marah, dan takut dengan hasil akurasi mencapai 65,93%. Imran, Daudpota, Kastrati, dan Batra[7] juga melakukan deteksi emosi dan polaritas *Cross-Cultural* mengenai *Tweets* yang terkait dengan COVID-19 dengan menggunakan analisis sentimen dan *Deep Learning LSTM*, yang menganalisis reaksi masyarakat dari berbagai negara terhadap COVID-19 juga sentimen mengenai tindakan yang akan dilakukan negara yang berbeda, dimana penelitian ini mencapai nilai akurasi sebesar 78%.

Yuxio Chen dkk [5] melakukan analisis sentimen positif dan negatif pada *Twitter*. Penelitian ini memberikan perhatian yang lebih besar ke emoji, dengan menggunakan metode *bi-sense emoji embedding* dan *Attention-Based LSTM Network*. Penelitian ini mencapai nilai akurasi mencapai 82,4%. Yequan, Minlie, Li, dan Xiaoyan [7] melakukan klasifikasi sentimen terhadap teks *review* dengan metode *Attention-based LSTM* dengan *Aspect embedding*. Penelitian ini efektif untuk memunculkan kalimat yang berbeda Ketika menyangkut aspek yang berbeda juga, dengan nilai akurasi paling tinggi mencapai 90,9%. Mapes, White, Medury, dan Dua [8] melakukan klasifikasi terhadap bahasa pemecah belah dan propaganda dari *Tweets* menggunakan *Attention Transformers* dengan BERT-based Language Models, dengan hasil yang cukup bagus dimana nilai akurasinya mencapai 60,1%. Amirmohammad Kazameini dkk [9] melakukan penelitian dimana mendeteksi sifat kepribadian seseorang, yang mengharapkan untuk mengembangkan model prediksi ini yang berkinerja tinggi dan efisien, yang nantinya dapat digunakan banyak orang tanpa perlu akses ke sumber daya komputer yang besar. Penelitian ini dilakukan

dengan menggunakan BERT *word embedding* dan SVM *classifier*, yang mencapai nilai akurasi 59,03%.

III. METODE PENELITIAN

Tujuan dari penelitian ini adalah mengklasifikasikan respons dan menganalisis hasil klasifikasi terhadap vaksinasi COVID-19 di Indonesia berdasar *Tweets* menggunakan metode BERT dan A-LSTM. Terdapat enam tahap dalam penelitian ini yaitu pengumpulan data, preprocessing data, anotasi data, tokenisasi kata, klasifikasi menggunakan A-LSTM, dan evaluasi. Tahap dalam penelitian ini dapat dilihat pada Gambar 1.

A. Pengumpulan Data

Data yang digunakan adalah *dataset* berupa kumpulan *tweet* berbahasa Indonesia dengan kata kunci “vaksin”, “vaksin covid”, dan “vaksin corona”. Pengumpulan data ini dilakukan dimulai tanggal 1 Desember 2020 hingga 31 Maret 2021 dan didapatkan 139445 data. Untuk contoh dari data yang telah diambil dapat dilihat pada Tabel 1.

B. Preprocessing Data

Tahap ini diperlukan supaya data menjadi lebih bersih dan komputer lebih mudah untuk memproses data. Proses pada tahap ini meliputi merubah huruf menjadi *lowercase* atau kecil semua, dan menghapus duplikat pada *tweet*, *mention*, *hashtag*, tautan, dan simbol yang tidak penting. Setelah dilakukan *preprocessing* pada seluruh data, jumlah data berkurang sehingga jumlah total data sebanyak 110626 data. Hasil *preprocessing* terhadap contoh data dapat dilihat di Tabel 2.

C. Anotasi Data

Dalam tahapan ini dilakukan proses anotasi pada *tweet*, yaitu memberi label pada setiap *tweet* sebelum data diproses. Tahap ini menggunakan *library* TextBlob, yang dapat menentukan polaritas pada *tweet*. Label yang diberikan adalah label positif, negatif, dan netral. *Tweet* dinyatakan memiliki label negatif jika nilai polaritas kurang dari nol, label netral jika nilai polaritas sama dengan nol, dan label positif jika nilai polaritas lebih dari nol. Setelah dilakukan anotasi pada seluruh 110626 data, didapat label negatif sebanyak 13854 data, label netral sebanyak 46712 data, dan label positif sebanyak 50060 data. Hasil dari proses anotasi pada contoh data dapat dilihat pada Tabel 3.

D. Tokenisasi Kata

Tokenisasi merupakan sebuah proses untuk membagi sejumlah teks baik kalimat maupun paragraf menjadi beberapa bagian tertentu, yang disebut dengan token. Tokenisasi pada penelitian ini menggunakan BERT yang diperkenalkan oleh [10], dimana dari tiap token akan didapatkan *embedding* token berupa angka yang menyesuaikan kosa kata dari model BERT yang digunakan. Penelitian ini menggunakan model BERT dikhususkan untuk Bahasa Indonesia yang sebelumnya telah dilatih dengan Wikipedia Bahasa Indonesia, dengan kosa kata sebesar 32000. Hasil proses tokenisasi pada contoh data dapat dilihat pada Tabel 4.

Keterangan [CLS] dan [SEP] merupakan token khusus pada BERT, dimana secara otomatis ditambahkan di awal dan

akhir setiap *input*. Keterangan [PAD] secara otomatis ditambahkan setelah token terakhir hingga panjangnya sama dengan data dengan jumlah token terbanyak. Terlihat pula beberapa token yang diawali dengan ## seperti ##nya. Hal ini disebabkan kata 'harusnya' tidak terdapat dalam kosa kata, sehingga model BERT secara otomatis memisahkan kata tersebut menjadi dua subkata yaitu 'harus' yang terdapat di kosa kata, dan yang terakhir '##nya'.

E. Klasifikasi

Proses klasifikasi pada penelitian ini menggunakan *Attention-based Long Short Term Memory* (A-LSTM). LSTM dikembangkan untuk menyelesaikan permasalahan *vanishing gradient* pada RNN (Recurrent Neural Network) dan mencapai kinerja yang baik[7]. Kemudian ditambahkan *attention mechanism* yang digunakan untuk memutuskan bagian teks yang berbeda, dimana biasanya terdapat beberapa kata yang informatif, dan terdapat juga kata yang tidak relevan, yang kemudian menggabungkan representasi kata-kata tersebut[11]. Model pada klasifikasi ini menggunakan empat *layer*, yaitu *layer embedding*, *layer LSTM*, *layer attention*, dan *layer dense*.

1) Layer Embedding

Layer embedding menjalankan proses *word embedding*, yang merupakan salah satu pendekatan dan metode dari pemrosesan bahasa alami, dimana proses ini memetakan kata atau frasa yang tersedia pada *vocabulary* ke dalam bentuk vektor dengan nilai bilangan real, dan telah terbukti dapat meningkatkan kinerja pada pemrosesan bahasa alami, salah satunya analisis sentimen[12]. Masukan pada *layer* ini berupa matriks berukuran ($k \times \text{input_length}$), dimana k merupakan jumlah *tweet* dalam satu proses, dan *input_length* yang merupakan panjang dari masukan. Sedangkan keluaran pada *layer* ini berupa matriks 3 dimensi berukuran ($k \times \text{input_length} \times \text{embedding_dim}$), dengan *embedding_dim* yang merupakan dimensi hasil dari *word embedding*. Terdapat pula parameter *input_dim* yang merupakan ukuran kosa kata yang dimiliki oleh model BERT, yang diinisiasikan pada model.

2) Layer LSTM

Layer LSTM menjalankan metode *Long Short-Term Memory*, dimana pada desain arsitekturnya, LSTM memiliki tiga *gate*/gerbang, yaitu *gate input*, *gate forget*, dan *gate output*. LSTM juga memiliki komponen bernama *cell layer*. Ketiga *gate* juga *cell layer* memiliki bobot dan bias, yang disimbolkan W dan b .

Misalkan h_{t-1} merupakan *hidden state* sebelumnya dengan ukuran ($n \times I$) dimana n merupakan jumlah *neuron* pada LSTM, dan x_t merupakan matriks kolom yang berukuran ($p \times I$) dimana p merupakan panjang dari hasil *word embedding*. Pada setiap komponen, h_{t-1} dan x_t memiliki bobot masing-masing, disimbolkan W_h dan W_x . W_h merupakan matriks persegi berukuran ($n \times n$), dan W_x merupakan matriks berukuran ($n \times p$). Sedangkan untuk bias b berukuran ($n \times I$).

Pada *gates forget*, ditentukan apakah informasi dari h_{t-1} dan x_t dapat dilanjutkan atau tidak, dengan hasilnya menggunakan persamaan:

$$f_t = \sigma(W_{hf} \cdot h_{t-1} + W_{xf} \cdot x_t + b_f) \quad (1)$$

dimana operator \cdot merupakan operator *dot product*, f_t adalah hasil dari *gate forget* pada waktu t , σ adalah fungsi *sigmoid*, kemudian W_{hf} merupakan bobot *forget* milik h_{t-1} , W_{xf} merupakan bobot *forget* milik x_t , dan b_f merupakan bias pada *forget gate*.

Berikutnya pada *gate input*, bertujuan untuk menentukan bagian informasi mana yang disimpan dalam *cell state*. Terdapat dua tahap dalam *gate* ini, yaitu terdiri dari fungsi *sigmoid* yang menentukan informasi yang diperbarui, dan fungsi *tanh* yang membuat kandidat informasi baru, dilambangkan \tilde{C}_t . Proses ini dilakukan dengan menggunakan persamaan :

$$i_t = \sigma(W_{hi} \cdot h_{t-1} + W_{xi} \cdot x_t + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_{hc} \cdot h_{t-1} + W_{xc} \cdot x_t + b_c) \quad (3)$$

dimana i_t adalah hasil dari *gate input* pada waktu t , W_i adalah bobot untuk *gate input*, b_i adalah bias untuk *gate input*, W_c adalah bobot untuk *cell layer*, dan b_c adalah bias untuk *cell layer*.

Berikutnya yaitu *cell layer*, dimana *cell state* C_t kemudian diperbarui dengan menggunakan persamaan:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

dimana operator $*$ merupakan *Hadamard product*, yang merupakan perkalian dari tiap elemen yang ada pada dua matriks berdimensi yang sama dan menghasilkan matriks dengan dimensi yang sama pula.

Yang terakhir merupakan *gate output*, dimana juga terdapat fungsi *sigmoid* yang memutuskan bagian *cell state* yang dijadikan *output*, dan fungsi *tanh* untuk membuat nilai menjadi diantara -1 dan 1. Tujuannya untuk membuat keputusan untuk *hidden state* h_t sebagai *output* dari LSTM. Proses ini dilakukan dengan menggunakan persamaan:

$$o_t = \sigma(W_{ho} \cdot h_{t-1} + W_{xo} \cdot x_t + b_o) \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

dimana o_t adalah *output* dari *gate output*, W_o adalah bobot dari *gate output*, dan b_o adalah bias untuk *gate output*.

Dalam penelitian ini, masukan pada *layer LSTM* ini merupakan keluaran dari *layer* sebelumnya. Sedangkan untuk keluaran pada *layer* ini merupakan matriks 3 dimensi berukuran ($k \times \text{input_length} \times \text{neuron}$), dimana *neuron* sendiri merupakan dimensi dari LSTM.

3) Layer Attention

Layer attention menjalankan *attention mechanism*, yang dapat memperhitungkan seluruh *hidden state*, menentukan nilai yang lebih penting, kemudian menjumlahkan seluruh nilai sehingga terbentuk *output* yang dapat meningkatkan kinerja model. Desain arsitektur pada penambahan *attention layer* ini seperti pada Gambar 3.

Misalkan $h = [h_1, h_2, \dots, h_t]$ merupakan *hidden state* hasil dari LSTM, dimana h_t *hidden state* ke- t yang berukuran ($n \times I$). Terdapat pula bobot dan bias untuk tambahan *attention layer* ini, disimbolkan W_a dan b_a . *Attention layer* akan menghasilkan *output* akhir z .

Beberapa persamaan yang dilakukan dalam *attention layer* ini sebagai berikut [14]:

$$M = \tanh(W_a \cdot h + b_a) \quad (7)$$

$$\alpha_t = \text{softmax}(M_t) = \frac{e^{M_t}}{\sum_{i=1}^t e^{M_i}} \quad (8)$$

$$z_t = h_t \alpha_t$$

$$z = \sum_{i=1}^t z_t$$

dimana $M = [M_1, M_2, \dots, M_t]$ berukuran $(1 \times t)$, W_a merupakan matriks bobot berukuran $(1 \times n)$, b_a merupakan bias berukuran $(1 \times t)$, α merupakan nilai dari *attention*, dan z berukuran $(n \times 1)$ adalah vektor konteks yang merupakan hasil terakhir dari *layer attention* dan LSTM, yang kemudian dilanjutkan ke *layer* berikutnya dan memprediksi kategori dari sebuah kalimat.

Masukan pada *attention layer* berupa keluaran dari *layer* sebelumnya, dan keluaran pada *layer* ini yaitu *output* berupa vektor konteks z berukuran $(k \times \text{neuron})$.

4) Layer Dense

Layer dense merupakan salah satu lapisan yang sering digunakan pada jaringan saraf, dimana fungsinya dapat merubah dimensi sebuah vektor dengan melakukan multiplikasi matriks dan vektor. Pada *layer* ini berlaku persamaan sebagai berikut:

$$\text{output} = \text{activation}((z \cdot W_d) + b_d) \quad (9)$$

dimana W_d dan b_d merupakan bobot dan bias pada *layer dense*. Penelitian ini menggunakan *activation softmax*, yang biasa digunakan sebagai fungsi aktivasi untuk permasalahan klasifikasi multi-kelas. Keluaran untuk *layer* ini merupakan matriks berukuran $(k \times 3)$, dengan 3 merupakan ketiga label yang telah dibuat di awal, sehingga *output* digunakan untuk memprediksi label positif, negatif, dan netral.

IV. HASIL DAN PEMBAHASAN

Sebelum dilakukan uji coba pada model yang telah dibuat, data terlebih dahulu dibagi menjadi tiga yaitu data latih, data validasi, dan data uji. Pertama data dibagi menjadi 5% data uji dan 95% data latih. Kemudian dari 95% data latih, diambil 10% data yang digunakan untuk validasi, sehingga perbandingan antara data latih dan data validasi adalah 90%:10%.

A. Uji Coba Model

Pada penelitian ini menggunakan 20 *epoch* dan *batch_size* sebesar 250. Dari proses *training* pada model dengan menggunakan data latih dan data validasi untuk percobaan pertama, didapatkan grafik akurasi dan *loss* tiap *epoch* terlihat pada Gambar 4 dan Gambar 5.

Pada grafik model *loss*, untuk garis *validation* terlihat mengalami penurunan nilai *loss* pada awalnya, namun kemudian nilai *loss* menjadi semakin besar. Hal ini biasa disebut *overfitting*. *Overfitting* merupakan kondisi dimana model telah melatih data latih terlalu baik, sehingga jika diberi data baru yaitu data validasi, model justru tidak bekerja dengan baik. Ada banyak penyebab terjadinya kondisi ini, seperti jumlah atau kualitas data yang kurang, maupun *input* variabel pada model itu sendiri yang kurang sesuai.

Untuk percobaan kedua, ditambahkan *dropout* pada pelatihan model. *Dropout* merupakan salah satu cara yang dapat menghindari *overfitting* dimana menghapus beberapa *neuron* secara acak untuk sementara dalam pelatihan model[13].

Setelah percobaan kedua dilakukan dengan jumlah *epoch* dan *batch_size* serupa, didapatkan grafik akurasi dan *loss* tiap *epoch* terlihat pada Gambar 6 dan Gambar 7.

Berdasarkan kedua grafik terlihat bahwa model sudah cukup baik dengan menunjukkan akurasi data latih dengan validasi tidak berbeda jauh, juga nilai *loss* yang cenderung stabil sehingga tidak mengalami *overfitting*.

Hasil dari model yang telah dibuat akan dibandingkan dengan model baru dengan metode LSTM, dimana pada model ini tidak menggunakan *layer attention*. Untuk hasil grafik nilai akurasi dan *loss* pada model LSTM dengan jumlah *epoch* dan *batch_size* yang sama terlihat pada Gambar 8 dan Gambar 9.

Pada Gambar 8 dan Gambar 9, selisih nilai tiap *epoch* antara data latih dan data validasi lebih kecil dibandingkan Gambar 6 dan Gambar 7. Sehingga, hasil dari proses *training data* pada model LSTM lebih bagus daripada model A-LSTM yang telah dibuat. Namun, nilai akurasi dan *loss* ini bukan merupakan nilai keseluruhan dari model. Untuk menguji keseluruhan model, dilakukan tahap evaluasi model.

B. Evaluasi Model

Tahap ini merupakan tahap dimana model diuji dengan menggunakan data uji. Tahap ini diperlukan supaya kinerja model terhadap proses klasifikasi dapat diketahui. Pada tahap evaluasi model, hasil klasifikasi dari data uji dikelompokkan menjadi empat, yaitu:

1. *True Positive* (TP) : Jumlah data positif yang diklasifikasikan ke positif oleh model
2. *True Negative* (TN) : Jumlah data negatif yang diklasifikasikan ke negatif oleh model
3. *False Positive* (FP) : Jumlah data negatif yang diklasifikasikan ke positif oleh model
4. *False Negative* (FN) : Jumlah data positif yang diklasifikasikan ke negatif oleh model

Pada penelitian ini evaluasi yang dilakukan adalah dengan melihat nilai akurasi, presisi, *recall*, dan F1. Nilai akurasi dapat digunakan untuk mengetahui seberapa sering data diklasifikasikan secara benar. Nilai presisi digunakan untuk mengukur ketepatan antara informasi yang diminta dan jawaban yang diberikan. Nilai *recall* digunakan untuk mengetahui tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi dengan benar. Nilai F1 dapat diartikan sebagai rata-rata dari nilai presisi dan nilai *recall*, yang dapat digunakan untuk membandingkan hasil klasifikasi. Perhitungan dari keempat nilai ini sebagai berikut.

$$\text{accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

$$\text{precision} = \frac{TP}{TP+FP} \quad (10)$$

$$\text{recall} = \frac{TP}{TP+FN} \quad (11)$$

$$F1 - \text{score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (12)$$

Dari keempat perhitungan diatas didapatkan keempat nilai untuk model A-LSTM dan model LSTM pada Tabel 5.

Setelah dilakukan proses anotasi pada seluruh 110626 data, dimana hasil label yang paling banyak diperoleh label positif dengan 50060 data, kemudian label netral 46712 data, dan label negatif 13854 data. Hal ini menunjukkan bahwa dari seluruh data *tweet* yang dikumpulkan, sebanyak 45,25% dari

seluruh respons bereaksi terhadap vaksinasi COVID-19 dengan positif, sehingga masyarakat dapat terpengaruh dengan respons positif tersebut dan bersedia menerima vaksin, juga ikut mensukseskan program vaksinasi COVID-19.

Berdasarkan keempat nilai pada Tabel 5, kedua model memiliki nilai yang serupa. Nilai presisi, *recall*, dan F1 pada label negatif lebih kecil dibandingkan dua label lainnya. Hal ini dapat disebabkan data yang kurang merata antara label negatif dan label lainnya. Untuk model A-LSTM memiliki enam kategori yang lebih unggul daripada LSTM biasa. Akurasi keseluruhan model A-LSTM juga lebih tinggi dibandingkan model LSTM.

Model A-LSTM sudah cukup bagus dalam melakukan tugas mengklasifikasikan antara tiga label yang telah dibuat. Dari acuan hasil akurasi pada penelitian sebelumnya, hasil nilai akurasi pada model sudah bagus dan tidak terlalu jauh dari penelitian yang sudah ada. Beberapa penelitian memiliki nilai akurasi yang lebih tinggi daripada hasil penelitian ini. Hal ini dapat disebabkan oleh data yang kurang banyak, kurang bersih, atau hasil anotasi oleh *library* TextBlob yang kurang akurat. Sehingga dibutuhkan data yang lebih banyak dan lebih bersih supaya dapat memberikan kinerja yang lebih baik.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Vaksinasi COVID-19 menimbulkan berbagai respons dari pengguna *Twitter* Indonesia. Setelah dilakukan proses anotasi pada 110626 data yang telah dikumpulkan, didapatkan 13854 respons negatif, 46712 respons netral, dan 50060 respons positif, sehingga masyarakat lebih cenderung merespons hal vaksinasi dengan positif. Konstruksi model A-LSTM mampu menentukan klasifikasi dari sebuah respons terhadap vaksinasi pada *tweet*. Seluruh data yang digunakan dibagi menjadi 5% data uji dan 95% data latih, dimana 10% dari data latih digunakan untuk validasi model. Respons pada *tweet* diproses menggunakan BERT untuk merubah kata menjadi angka, dan dilatih menggunakan empat *layer* yaitu *layer embedding*, *layer LSTM*, *layer Attention*, dan *layer Dense*, sehingga didapatkan model yang mampu mengklasifikasi respons berdasarkan data latih yang telah dipelajari.

Nilai kinerja yang didapat dari pengujian model dengan menggunakan data uji sebanyak 5532 data memiliki akurasi sebesar 82%. Model ini cukup baik dan memiliki nilai akurasi yang lebih tinggi jika dibandingkan dengan LSTM biasa. Walaupun begitu, jarak nilai *loss* pada grafik *train* dan *validation* masih cukup jauh, sehingga masih memungkinkan atas terjadinya *overfitting* jika dilakukan proses *training* lebih lanjut.

B. Saran

Membersihkan data dan melakukan anotasi pada data dengan lebih akurat.

Mengembangkan model untuk melihat pengaruh yang lebih baik pada model.

DAFTAR PUSTAKA

- [1] H. Tankovska, "Number of social network users in selected countries in 2020 and 2025,","
- [2] Kementerian Kesehatan, ITAGI, UNICEF, and WHO, "Survei Penerimaan Vaksin COVID-19 di Indonesia," 2020.
- [3] R. R. Aryal and A. Bhattarai, "Sentiment analysis on Covid-19 vaccination tweets using Naïve Bayes and LSTM," *Advances in Engineering and Technology: An International Journal*, vol. 1, no. 1, pp. 57–70, 2021, doi: 10.3126/AET.V1I1.39660.
- [4] A. S. Imran, S. M. Daudpota, Z. Kastrati, and R. Batra, "Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets," *IEEE Access*, vol. 8, pp. 181074–181090, 2020, doi: 10.1109/ACCESS.2020.3027350.
- [5] Y. Chen, Q. You, J. Yuan, and J. Luo, "Twitter sentiment analysis via bi-sense emoji embedding and attention-based LSTM," in *MM 2018 - Proceedings of the 2018 ACM Multimedia Conference*, 2018, pp. 117–125. doi: 10.1145/3240508.3240533.
- [6] E. Junianto and R. Rachman, "Penerapan metode Naïve Bayes Classifier untuk mendeteksi emosi pada komentar media sosial," *Jurnal Responsif: Riset Sains dan Informatika*, vol. 2, no. 1, pp. 1–8, 2020, doi: 10.51977/JTI.V2I1.156.
- [7] Y. Wang, M. Huang, L. Zhao, and X. Zhu, "Attention-based LSTM for aspect-level sentiment classification," in *EMNLP 2016 - Conference on Empirical Methods in Natural Language Processing, Proceedings*, 2016, pp. 606–615. doi: 10.18653/V1/D16-1058.
- [8] N. Mapes, A. White, R. Medury, and S. Dua, "Divisive Language and Propaganda Detection using Multi-head Attention Transformers with Deep Learning BERT-based Language Models for Binary Classification," in *Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda*, 2019, pp. 103–106. doi: 10.18653/V1/D19-5014.
- [9] A. Kazemeini, S. Fatehi, Y. Mehta, S. Eetemadi, and E. Cambria, "Personality trait detection using bagged SVM over BERT word embedding ensemble," Oct. 2020, doi: 10.6084/M9.FIGSHARE.13012421.V1.
- [10] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 2018, vol. 1, pp. 4171–4186. doi: 10.48550/arxiv.1810.04805.
- [11] C. W. Chen, S. P. Tseng, T. W. Kuan, and J. F. Wang, "Outpatient text classification using attention-based Bidirectional LSTM for robot-assisted servicing in hospital," *Information*, vol. 11, no. 2, p. 106, 2020, doi: 10.3390/INFO11020106.
- [12] D. Ayata, M. Saraclar, and A. Özgür, "BUSEM at SemEval-2017 Task 4A sentiment analysis with word embedding and long short term memory RNN approaches," in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, 2017, pp. 777–783. doi: 10.18653/V1/S17-2131.
- [13] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 56, pp. 1929–1958, 2014.