Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 21, No. 3, Juli 2022, pp. 607∼618
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: 10.30812/matrik.v21i3.1402

Sentimen Ulasan Destinasi Wisata Pulau Bali Menggunakan *Bidirectional Long Short Term Memory*

Sentiment of Balis Touristic Destination Reviews Using Bidirectional Long Short Term Memory

Dwi Intan Afidah[1], Dairoh[2], Sharfina Febbi Handayani[3], Riszki Wijayatun Pratiwi[4], Susi Nurindah Sari[5]
Politeknik Harapan Bersama, Indonesia

Informasi Artikel

Genesis Artikel:

Diterima, 17 Agustus 2021
Direvisi, 24 Februari 2022
Disetujui, 27 Mei 2022

Kata Kunci:

Analisis Sentimen
Bidirectional Long Short Term Memory
Word2Vec

ABSTRAK

Pemerintah dan pelaku industri pariwisata mengalami permasalahan dalam menentukan prioritas pengembangan suatu destinasi wisata. Karena itu, diperlukan identifikasi objek wisata yang diminati namun banyak mendapat ulasan buruk melalui ulasan dari masyarakat yang tersebar di internet. Penelitian ini bertujuan melakukan analisis sentimen terhadap ulasan objek wisata di Pulau Bali menggunakan *Bi-LSTM* dan *Word2Vec*, sehingga diperoleh model terbaik yang dapat digunakan untuk mengidentifikasi objek wisata potensial namun mendapat ulasan buruk. *Bi-LSTM* merupakan *deep learning* yang menawarkan akurasi yang lebih baik daripada LSTM biasa. Sedangkan *Word2Vec* merupakan *pretraining* yang dipilih karena dapat menangkap makna semantik teks. Penelitian ini menggunakan data ulasan objek wisata di Pulau Bali yang berasal dari situs tripadvisor.com. Penelitian dimulai dari pengumpulan data, perancangan alur program, *preprocessing*, *pretraining Word2Vec*, pembagian data uji dan data latih, pelatihan dan pengujian, serta evaluasi penentuan model terbaik. Akurasi terbaik dihasilkan oleh kombisasi *Word2Vec* terdiri dari CBOW, *Hierarchical Softmax*, dimensi 200, *Bi-LSTM* dengan *dropout* sebesar 0,5 dan *learning rate* sebesar 0,0001. Kombinasi tersebut menghasilkan akurasi tertinggi dari keseluruhan 108 kombinasi yaitu sebesar 96,86%, *precission* sebesar 96,53%, *Recall* sebesar 96,31%, F1 *Measure* sebesar 96,41%. Akurasi yang baik tersebut membuktikan bahwa kombinasi parameter *Bi-LSTM* dan *Word2Vec* cocok digunakan untuk analisis sentimen ulasan objek wisata di Pulau Bali.

Keywords:

Sentiment Analysis
Bidirectional Long Short Term Memory
Word2Vec

ABSTRACT

Government and tourism industry having problems in determining priorities for the development of a tourist destination. Therefore, is necessary to identify the tourist attractions which are in great demand but received a lot of bad reviews through reviews from the public scattered on the internet. This study aims to sentiment analysis on reviews of Bali Islands tourist attraction using the Bi-LSTM and Word2Vec, to obtain the best model which can be used to identify potential tourist attractions but received a lot of bad reviews. Bi-LSTM is a deep learning method that show better accuracy than the regular LSTM. Meanwhile, Word2Vec as one of the pretraining methods was chosen because it can capture the semantic meaning of the text. This research used data form reviews of Bali tourist attraction on tripadvisor.com site. The research flow included data collection, program flow design, preprocessing, Word2Vec pretraining, data sharing into test and training data, training and testing processes, and evaluation of test results to determine the best model. The best accuracy is produced by a combination namely Word2Vec consist of the CBOW, Hierarchical Softmax, 200 dimensions, Bi-LSTM specifically dropout 0.5, and learning rate of 0.0001. This combination produces the highest accuracy of 108 combinations, which is 96,86%, precision is 96.53%, Recall is 96.31%, F1 Measure is 96.41%. This good accuracy proves that the combination of Bi-LSTM and Word2Vec parameters is suitable for sentiment analysis of Balis touristic destination reviews.

Penulis Korespondensi:

Dwi Intan Afidah,
Program Studi DIV Teknik Informatika,
Politeknik Harapan Bersama,
Email: dwiintanafidah@poltektegal.ac.id

1. PENDAHULUAN

Pariwisata memberikan sumbangsih terbesar terhadap peningkatan devisa pada setiap negara. Indonesia termasuk salah satu negara yang mengandalkan pariwisata sebagai sumber utama devisa. Pariwisata di Pulau Bali merupakan salah satu pariwisata di Indonesia yang memberikan kontribusi terbesar dalam mendatangkan devisa negara. Pariwisata di Pulau Bali selain mendatangkan devisa, juga menjadi sumber pendapatan daerah [1]. Sebagai tujuan utama bagi wisatawan asing maupun wisatawan lokal, Pulau Bali perlu mendapatkan perhatian khusus dari pemerintah. Pengembangan pariwisata di Pulau Bali memegang peranan penting dalam persaingan ekonomi global karena Bali menjadi citra pariwisata Indonesia di kancah internasional. Selain itu, pengelolaan pariwisata yang tepat di Pulau Bali terutama pada objek wisata baru yang potensial akan meningkatkan jumlah kedatangan wisatawan, sehingga dapat meningkatkan devisa negara dan pendapatan daerah. Akan tetapi, pemerintah dan pelaku industri pariwisata mengalami permasalahan dalam menentukan prioritas pengembangan suatu destinasi wisata. Karena itu, diperlukan identifikasi objek wisata yang diminati namun banyak mendapat ulasan buruk melalui ulasan dari masyarakat yang tersebar di internet.

Masyarakat pada umumnya akan mencari informasi sebagai bahan pertimbangan sebelum memilih tujuan wisata. Saat ini masyarakat akan saling bertukar informasi mengenai objek wisata melalui media dari internet, seperti media sosial, *travel blog*, dan situs ulasan objek wisata. Ulasan objek wisata yang dijadikan referensi wisatawan biasanya berasal dari opini, usulan, atau argumen wisatawan lain yang sudah mengunjungi suatu objek wisata. Ulasan objek wisata tersebut dapat bersifat positif dan negatif [2]. Ulasan wisatawan ini menjadi penting karena dapat menjadi alat bantu baik bagi pemerintah dalam pengambilan keputusan untuk program pengembangan pariwisata maupun bagi wisatawan. Penentuan sentimen pada ulasan objek wisata Pulau Bali dengan bantuan manusia memiliki kekurangan karena memerlukan ahli dan waktu pengolahan data lama, oleh sebab itu diperlukan algoritma dan program yang mampu melakukan analisis sentimen terhadap ulasan objek wisata Pulau Bali.

Analisis sentimen merupakan salah satu teknik *Natural Language Processing* (NLP) yang menganalisis pendapat, sikap, dan emosi terhadap suatu entitas yang berupa teks. Analisis sentimen diperlukan sebagai bahan evaluasi yang selanjutnya menjadi dasar dalam pengambilan keputusan [3]. Kesulitan dalam analisis sentimen biasanya terjadi karena terlalu banyaknya data. *Deep learning* dapat menyelesaikan masalah banyaknya data dengan menunjukkan kinerja yang lebih baik dalam analisis sentimen dibandingkan dengan *machine learning* klasik. Berlawanan dengan machine learning klasik yang membutuhkan fitur seleksi, deep learning tidak membutuhkan fitur seleksi [4]. Pada suatu penelitian SVM (Support Vector Machine) terbukti memiliki performa lebih baik dibandingkan model *machine learning* klasik lainnya [5]. Akan tetapi, perbandingan metode SVM sebagai model *machine learning* klasik terbaik dengan LSTM (*Long Short Term Memory*) sebagai model *deep learning* membuktikan bahwa LSTM memberikan kinerja lebih baik daripada SVM [6].

Long Short Term Memory (LSTM) merupakan salah satu pengembangan *Recurrent Neural network* (RNN) untuk mengatasi masalah difusi gradien. Penelitian [7] membandingkan metode RNN dan LSTM analisis sentimen teks yang panjang dengan metode pretraining Word2Vec. Data analisis yang digunakan berupa komentar JD.COM (salah satu online shop di Cina), Ctrip Travel dari Cina, dan ulasan film. Hasil dari penelitian ini menunjukkan bahwa LSTM memiliki kinerja yang lebih baik dari RNN konvensional dalam melakukan klasifikasi teks pada semua sumber data pada penelitian ini [7]. Kombinasi dari dua metode *deep learning* atau lebih dapat dilakukan pada suatu analisis sentimen. Penelitian [8] menggunakan dua metode *deep learning* sekaligus yakni *Long Short Term Memory-Convolutional Nerual Network*(LSTM-CNN) serta menggunakan *Word2Vec* sebagai metode *pretraining*. LSTM-CNN dikombinasikan untuk menjadi solusi dari kelemahan masing-masing yang dimiliki LSTM tunggal dan CNN tunggal. Penelitian ini menggunakan dataset berupa ulasan objek wisata Pulau Bali. Hasil sentimen analisis penelitian ini menunjukkan adanya perbaikan akurasi dari metode LSTM-CNN dibandingkan metode LSTM tunggal [8, 9].

Bidirectional Long Short Term Memory (*Bi-LSTM*) merupakan kombinasi metode *deep learning* yang terdiri dari dua buah *layer* LSTM. Jadi *Bidirectional Long Short Term Memory* (*Bi-LSTM*) adalah pengembangan dari yang memungkinkan pelatihan tambahan dengan melintasi data masukan dua kali yaitu, dari kiri ke kanan, dan dari kanan ke kiri. Perbandingan Bi-LSTM dan LSTM pada data teks berbagai bahasa menunjukkan bahwa pelatihan data tambahan dari Bi-LSTM menawarkan akurasi yang lebih baik daripada metode LSTM biasa [10–13]. Penelitian [10] melakukan klasifikasi emosi pada teks berbahasa Indonesia menggunakan metode *Bi-LSTM* dan *pretraining Glove Word*. Dataset yang digunakan berupa lirik lagu berbahasa Indonesia. Emosi dari lirik lagu dikategorikan menjadi: marah, senang, sedih, dan tenang. Penelitian ini menunjukkan bahwa akurasi dari model terbaik *Bi-LSTM* dan *pretraining Glove Word* sebesar 91,08% [10]. Adapun penelitian [14] membandingkan metode LSTM, CNN, RNN, *Naives Bayesian*, dan *Bi-LSTM* untuk analisis sentimen pada ulasan film berbahasa Mandarin. Penelitian ini membuktikan bahwa Bi-LSTM menghasilkan performa yang lebih baik dibandingkan metode lainnya [14].

Masalah lain yang muncul dalam analisis sentimen adalah penentuan metode *pretraining* yang tepat agar diperoleh model yang lebih akurat. *Word2Vec* sebagai salah satu metode *pretraining* dipilih karena dapat menangkap makna semantik teks dengan baik dan setiap kata yang berhubungan dicirikan dengan vektor yang cenderung mirip [15, 16]. *Word2Vec* diusulkan oleh sebagai jaringan syaraf yang memproses data teks. *Word2Vec* mencakup dua model pembelajaran yaitu *Continuous Bag of Words* (CBOW) dan *Skip-gram*. *Word2Vec* dapat dimanfaatkan untuk mengelompokkan kata-kata yang serupa [17].

Berdasarkan masalah yang diuraikan, penelitian ini bertujuan untuk memperoleh model terbaik dari metode *Bi-LSTM* dan *Word2Vec* terhadap analisis sentimen teks ulasan objek wisata di Pulau Bali. Analisis sentimen ulasan objek wisata di Pulau Bali dibutuhkan sebagai bahan pertimbangan pemerintah dalam mengembangkan objek wisata potensial yang kurang diminati. Sedangkan bagi wisatawan, analisis sentimen dapat menjadi referensi dalam mempertimbangan kunjungan wisata ke Pulau Bali.

Organisasi penulisan artikel ini terdiri dari beberapa bagian. Bagian metode penelitian menjelaskan langkah-langkah penelitian yang dimulai dari pengumpulan data, perancangan alur program, *preprocessing*, *pretraining Word2Vec*, pembagian data uji dan data latih, pelatihan dan pengujian, serta evaluasi penentuan model terbaik. Bagian hasil dan analisis menjelaskan hasil dari tiap langkah yang terdapat pada metode penelitian. Bagian terakhir merupakan sub bab kesimpulan yang berisi kesimpulan dari hasil penelitian dan saran untuk penelitian selanjutnya.

2. METODE PENELITIAN

Penelitian ini bertujuan melakukan analisis sentimen terhadap ulasan objek wisata di Pulau Bali menggunakan metode *Bi-LSTM* dan *Word2Vec*, sehingga diperoleh model terbaik dari kombinasi parameter *Bi-LSTM* dan *Word2Vec*. Model analisis sentimen ini diperlukan bagi pemerintah dan pelaku industri pariwisata sebagai solusi permasalahan dalam menentukan prioritas pengembangan suatu destinasi wisata melalui identifikasi objek wisata yang diminati namun banyak mendapat ulasan buruk diperlukan. Adapun metode *Bi-LSTM* merupakan metode *deep learning* yang menawarkan akurasi yang lebih baik daripada metode LSTM biasa. Sedangkan *Word2Vec* merupakan metode pretraining yang dipilih karena dapat menangkap makna semantik teks.

2.1. Data Penelitian

Pada penelitian ini data yang digunakan merupakan data yang bersumber dari penelitian [8] berupa sekumpulan teks ulasan objek wisata Pulau Bali berbahasa Indonesia yang diambil dari situs tripadvisor.com. Data tersebut diperoleh menggunakan metode *web scraping* pada Bulan Februari 2020.

2.2. Prosedur Penelitian

Prosedur penelitian ini terdiri dari pengumpulan data, perancangan alur program, persiapan dataset terdiri dari *preprocessing* dan *pretraining Word2Vec*, pembagian data, pembentukan model, serta evaluasi. Secara keseluruhan prosedur penelitian yang dilakukan pada penelitian ini seperti ditunjukkan Gambar 1.

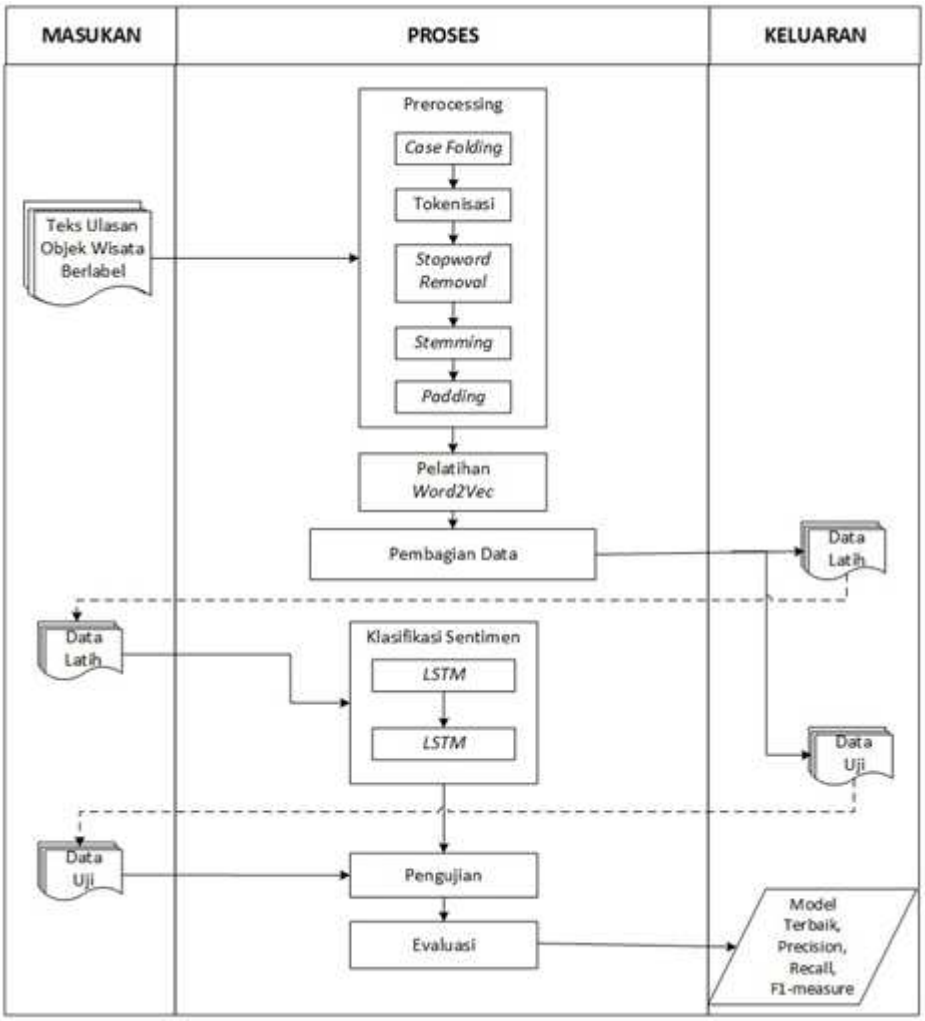
Gambar 1. Prosedur Penelitian

2.3. Pengumpulan Data

Data yang digunakan pada penelitian ini merupakan ulasan objek wisata di Pulau Bali berbahasa Indonesia pada situs tripadvisor.com yang berasal dari penelitian sebelumnya yakni penelitian [11]. Data dibagi secara seimbang yaitu, 5.000 data ulasan bersifat positif dan 5.000 data ulasan bersifat negatif. Data selanjutnya disimpan pada format json untuk dapat digunakan untuk proses pembentukan model klasifikasi analisis sentiment.

2.4. Perancangan Alur Program

Perancangan alur program digunakan untuk menentukan langkah-langkah apa saja perlu dilakukan dalam proses pembentukan model terbaik. Hasil perancangan alur program ini selanjutnya diterjemahkan ke dalam kode program agar setiap bagian dari alur program dapat memproses data sesuai tujuan. Alur program terdiri dari 2 alur yaitu alur untuk pelatihan model dan alur untuk pengujian ulasan. Proses pelatihan dilakukan untuk membentuk model, sedangkan proses pengujian dilakukan untuk memvalidasi model yang sebelumnya terbentuk. Semua data ulasan objek wisata akan diproses lebih dulu oleh *layer preprocessing* agar menjadi data rapi dan tidak terjadi redudansi. Kemudian data yang berupa teks akan diubah menjadi vektor supaya dapat dibaca oleh *deep learning*. Data selanjutnya dibagi menjadi data latih dan data uji. Hasil model dari proses pelatihan akan dilakukan pengujian menggunakan data uji. Hasil yang didapatkan pengujian kemudian dievaluasi sehingga diperolehnya model terbaik. Adapun gambaran alur program ditunjukkan pada Gambar 2.



Gambar 2. Alur Program

2.5. *Preprocessing* Data

Preprocessing teks atau pra pengolahan teks bertujuan menghilangkan redudansi data dan merapikan data agar mudah digunakan pada proses selanjutnya. Berikut ini merupakan tahapan pada *preprocessing* antara lain:

1. *Case folding*
   Case folding merupakan proses untuk mengubah semua karakter huruf pada sebuah kalimat menjadi huruf kecil atau huruf besar. Case folding yang dilakukan pada penelitian ini yaitu mengubah seluruh dataset menjadi huruf kecil.
2. Tokenisasi
   Tokenisasi merupakan proses untuk memecah dokumen teks menjadi token. Tokenisasi memiliki kemampuan untuk memecah dokumen menjadi kata, frasa, simbol atau elemen lain yang memiliki makna. Optimalisasi token dapat dilakukan dengan cara menghilangkan karakter-karakter ilegal pada dokumen seperti tanda baca, simbol, angka, html, dan mention.
3. *Stopword Removal*
   Stopword Removal merupakan tahap pengambilan kata-kata penting dan membuang kata-kata yang dianggap tidak penting. Stopword removal bertujuan untuk menghilangkan kata-kata yang sering muncul namun tidak memiliki kontribusi dalam proses analisis data.
4. *Padding*
   Proses pembelajaran yang dilakukan oleh neural network memerlukan masukan data dengan panjang yang sama. Padding merupakan proses yang dilakukan untuk membuat data input mempunyai panjang yang sama dengan cara menambahkan kata <pad>.

2.6. *Pretraining* Metode *Word2Vec*

Proses pelatihan model *Word2Vec* dimulai dengan menentukan data input dan data konteks, selanjutnya proses pelatihan akan menghasilkan data dalam representasi vektor. Data yang digunakan pada pelatihan model *Word2Vec* merupakan data ulasan objek wisata yang telah dilakukan *preprocessing* teks. *Array* hasil dari pelatihan *Word2Vec* kemudian disimpan dalam file berekstensi model. Proses model *Word2Vec* dijelaskan seperti pada Gambar 3.

Gambar 3. Pretraining Metode Word2Vec

Word2Vec memiliki beberapa parameter dalam proses pelatihan antara lain, arsitektur, metode evaluasi, dan dimensi dimana masing-masing parameter memiliki kategori. Tipe dari masing-masing parameter *Word2Vec* yang akan diujikan pada penelitian ini antara lain:

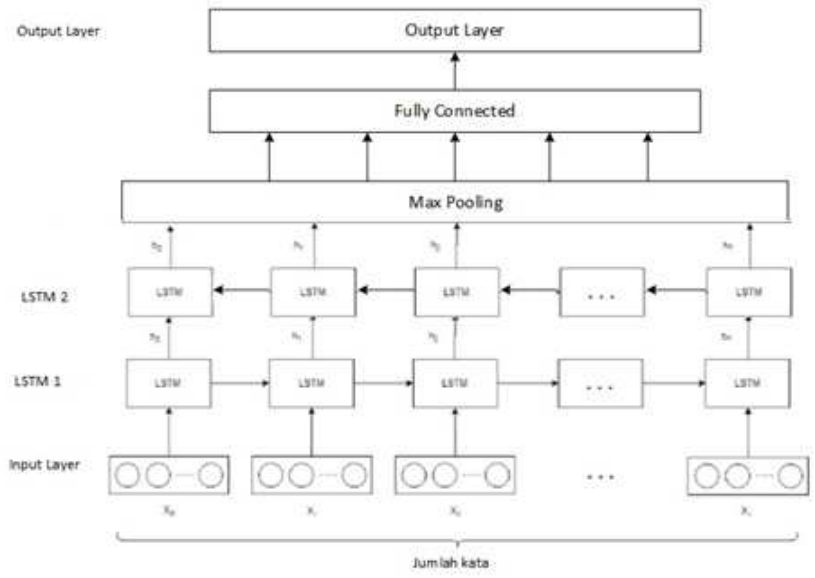
1. Arsitektur *Word2Vec*: CBOW (*Continuous Bag of Words*), dan *Skip-gram-gram*
2. Metode evaluasi: *Hierarchical Softmax*, dan *Negative Sampling*
3. Ukuran dimensi: 100, 200, dan 300

2.7. Pembagian data

Pembagian data dilakukan dengan membagi data menjadi 80% sebagai data latih dan 20% sebagai data uji. Pembagian data ini berlandaskan *scaling law* yang ditemukan oleh Guyon pada tahun 1997 [18]. Penelitian ini membahas pemisahan data latih dan data uji terbaik untuk masalah tertentu: mencegah pelatihan jaringan saraf yang berlebihan. Penelitian menemukan bahwa fraksi pola yang disediakan untuk data uji harus berbanding terbalik dengan akar kuadrat dari jumlah parameter yang dapat disesuaikan secara bebas. Pada intinya pemisahan ditentukan oleh berapa banyak fitur unik yang ada dalam kumpulan data (tidak termasuk target) dan bukan juga jumlah pengamatan.

2.8. Pembentukan Model

1. Rancangan Struktur Model
   Rancangan struktur model digambarkan pada Gambar 4.



Gambar 4. Rancangan Arsitektur Bi-LSTM dan Word2Vec

Gambar 4 input dari arsitektur model *Bi-LSTM* adalah *array* hasil *preprocessing* teks dengan ukuran Jumlah Kata. Masing-masing kata selanjutnya masuk ke dalam *embedding layer* untuk dicarikan representasi data vektornya pada *array* hasil *Word2Vec*. *Array* hasil *Word2Vec* kemudian diproses ke dalam *Bi-LSTM*. Output *Bi-LSTM* akan masuk pada *pooling layer*. Hasil dari dari *pooling layer* masuk ke *flatten layer*, setelah itu masuk melalui *dense layer* dengan fungsi aktivasi *sigmoid* yang kemudian akan mengeluarkan *output* dengan ukuran 1x1.

2. Pelatihan
   Proses pelatihan bertujuan untuk melakukan pelatihan menggunakan data yang telah diperoleh untuk mendapatkan hasil pemodelan yang terbaik. Parameter yang di inisialisasi pada model *Bi-STM* yaitu *dropout*, *pooling*, aktivasi *output*, *optimizer*, *learning rate*, serta parameter *Word2Vec* yang sudah dilatih sebelumnya. Seluruh kombinasi pada penelitian ini menggunakan *optimizer* Adam dan jumlah *output* node 1. Parameter *dropout*, *pooling*, dan *learning rate* dapat diubah sesuai dengan kombinasi yaitu *dropout* sebesar 0,2; 0,5; 0,7, *pooling* berupa *max pooling* atau *average pooling*, dan dengan nilai *learning rate* sebesar 0,001 dan 0,0001. Oleh karena itu, pelatihan satu per satu dilakukan pada setiap kombinasi parameter *Bi-LSTM*, dan parameter *Word2Vec* yang sudah dilatih sebelumnya.
3. Pengujian
   Proses pengujian dilakukan pada proses pembentukan model setelah dilakukan proses pelatihan. Tujuan dari proses pengujian yaitu untuk melakukan validasi dari hasil yang sudah dilatih melalui proses pelatihan pada setiap parameter yang diujikan. Pengujian dilakukan dengan mengevaluasi pengukuran akurasi terhadap data uji pada seluruh kombinasi. Proses pengujian membutuhkan data uji dan menggunakan semua model dari hasil pelatihan pada setiap kombinasi parameter.

2.9. Evaluasi

Hasil kinerja model klasifikasi penelitian ini diukur menggunakan *confusion matrix*. *Confusion matrix* merupakan *tools* untuk mengukur performa klasifikasi dokumen terhadap satu kelas atau lebih. Pada Tabel 1 menggambarkan contoh untuk prediksi dua kelas dengan menggunakan *confusion matrix*.

Tabel 1. Pembagian data untuk *Training* dan *Testing*

Parameter		Prediksi	
		Negatif(-)	**Positif(+)**
Aktual	Negatif (-)	TN	FN
	Positif (+)	FP	TP

Pada Tabel 1 diatas keterangan yang diperoleh bahwa *True Negative* jika hasil prediksi negatif dan data aktualnya negatif, True Positif jika hasil prediksi positif dan data aktualnya positif, kemudian False Negatif jika hasil prediksi negatif dan data aktualnya positif dan *False* Positif jika hasil prediksi positif dan data aktualnya negatif. Terdapat persamaan ditetapkan pada matriks dua kelas yang memiliki persamaan seperti pada Persamaan (1, 2, 3 dan 4).

$$Akurasi = \frac{(TN + TP)}{(TN + FN + TP + FP)} \tag{1}$$

$$Recall = \frac{(TP)}{(TP + FN)} \tag{2}$$

$$Precision = \frac{(TP)}{(TP + FP)} \tag{3}$$

$$Precision = \frac{(2 * Recall * Precision)}{(Recall + Precision)} \tag{4}$$

3. HASIL DAN ANALISIS

3.1. Hasil

Penelitian ini menghasilkan suatu model terbaik dari *Bidirectional Long Short Term Memory* untuk Analisis Sentimen terhadap Ulasan Destinasi Wisata Pulau Bali. Hasil dari penelitian ini dapat membantu bagi peneliti yang ingin mengembangkan aplikasi yang mampu melakukan sentimen analisis terhadap ulasan objek wisata.

1. Hasil pengumpulan data
   Sampel dari pengumpulan data dapat dilihat pada Tabel 2.

Tabel 2. Sampel Hasil Pengumpulan Data

No	Ulasan	Sentimen
1	Terkesan akan keindahan alamnya, pantai yang bersih. Tempatnya enak, makanannya enak. Cocok buat acara keluarga	1
2	Nggak begitu berkesan, tempat ini semakin nggak menarik. Air kotor, wahana permainan tidak aman, lokasi jauh.	0
3	Air terjun, pemandangan bagus dan penduduk yang ramah lokasi bersih. bawa pakain ganti untuk setelah berenang.	1
4	Jangan berekspektasi lebih di tempat ini, tempatnya spti pasar tradisional biasa, dgn desek desekan. memang dari harga bisa lebih miring tergantung kemampuan tawar menawar kita . . . sekedar saran lebih nyaman belanja di erlangga 2 atau krisna . . . saya sedikit menyesal ke sini	0

2. Hasil *preprocessing*
   Hasil dari masing-masing tahapan *prepocessing* dijelaskan sebagai berikut: Tabel 3 menunjukkan hasil *proses case folding*, Tabel 4 menunjukkan hasil proses tokenisasi, Tabel 5 menunjukkan hasil proses *stopword removal*, dan Tabel 6 menunjukkan hasil proses *padding*.

Tabel 3. Sampel Data Sebelum dan Setelah *Case Folding*

No	Data Sebelum *Case Folding*	Data Setelah *Case Folding*
1	Terkesan akan keindahan alamnya, pantai yang bersih. Tempatnya enak, makanannya enak. Cocok buat acara keluarga . . .	Terkesan akan keindahan alamnya, pantai yang bersih. tempatnya enak, makanannya enak. cocok buat acara keluarga.
2	Nggak begitu berkesan, tempat ini semakin nggak menarik. Air kotor, wahana permainan tidak aman, lokasi jauh	Nggak begitu berkesan, tempat ini semakin nggak menarik. air kotor, wahana permainan tidak aman, lokasi jauh

Tabel 4. Sampel Data Sebelum dan Setelah Tokenisasi

No	Data Sebelum Tokenisasi	Data Setelah Tokenisasi
1	Terkesan akan keindahan alamnya, pantai yang bersih. Tempatnya enak, makanannya enak. Cocok buat acara keluarga . . .	[terkesan, akan, keindahan, alamnya, pantai, yang, bersih, tempatnya, enak, makanannya, enak, cocok, buat, acara, keluarga]
2	Nggak begitu berkesan, tempat ini semakin nggak menarik. Air kotor, wahana permainan tidak aman, lokasi jauh	[nggak, begitu, berkesan, tempat, ini, semakin, nggak, menarik, air, kotor, wahana, permainan, tidak, aman, lokasi, jauh]

Tabel 5. Sampel Data Sebelum dan Setelah *Stopword Removal*

No	Data Sebelum *Stopword Removal*	Data Setelah *Stopword Removal*
1	[terkesan, akan, keindahan, alamnya, pantai, yang, bersih, tempatnya, enak, makanannya, enak, cocok, buat, acara, keluarga]	[terkesan, keindahan, alamnya, pantai, bersih, tempatnya, enak, makanannya, enak, cocok, buat, acara, keluarga]
2	[nggak, begitu, berkesan, tempat, ini, semakin, nggak, menarik, air, kotor, wahana, permainan, tidak, aman, lokasi, jauh]	[nggak, begitu, berkesan, tempat, semakin, nggak, menarik, air, kotor, wahana, permainan, tidak, aman, lokasi, jauh]

Tabel 6. Sampel Data Sebelum dan Setelah *Padding*

No	Data Sebelum *Padding*	Data Setelah Padding
1	[terkesan, akan, keindahan, alamnya, pantai, yang, bersih, tempatnya, enak, makanannya, enak, cocok, buat, acara, keluarga]	[kesan, indah, alam, pantai, bersih, tempat, enak, makan, enak, cocok, buat, acara, keluarga, ¡pad¿, ¡pad¿]
2	[nggak, begitu, berkesan, tempat, ini, semakin, nggak, menarik, air, kotor, wahana, permainan, tidak, aman, lokasi, jauh]	[nggak, begitu, kesan, tempat, semakin, nggak, tarik, air, kotor, wahana, main, tidak, aman, lokasi, jauh]

3. Hasil *Pretraining* Menggunakan *Word2Vec*
   Hasil dari *Pretraining* menggunakan metode *Word2Vec* adalah setiap kata memiliki terjemahan yang berupa vektor. Gambar 5 menunjukkan hasil vektor yang mewakili kata bagus yang dengan dimensi vektor sebesar 300.

```
1 word_vectors.wv['bagus']

/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: DeprecationWarning:
  """Entry point for launching an IPython kernel.
array([-0.13993983,  0.11706576,  0.0531881 ,  0.05269541, -0.10505376,
       -0.13594835,  0.04493795,  0.09740925,  0.09784161,  0.00330711,
        0.13818163,  0.01747776, -0.20778643, -0.03620799,  0.10143393,
        0.09320962,  0.06552859,  0.02472059, -0.00219108, -0.08866776,
        0.05293579, -0.06520167,  0.01089241, -0.0016338 , -0.05291765,
        0.00755603,  0.09149327, -0.01619183,  0.13862973,  0.06191736,
       -0.18147431, -0.04698475, -0.09193318,  0.00188438, -0.01461697,
       -0.1396031 ,  0.00707596,  0.14138715,  0.01063238, -0.11313259,
       -0.02921242,  0.10558522,  0.02787341, -0.16666207, -0.0307539 ,
        0.00140733, -0.05515899,  0.1758352 , -0.2335766 , -0.04746294,
       -0.26940772,  0.09067795,  0.0509332 ,  0.21438947, -0.18411247,
       -0.01498233,  0.02144537,  0.02699619,  0.06118437, -0.00711128,
        0.06009525, -0.04301694,  0.11626563, -0.08653332, -0.08962329,
       -0.03919519,  0.13904612, -0.07939364, -0.12208824,  0.01339524,
       -0.07694787, -0.06731313,  0.00713928, -0.05568055, -0.14392239,
       -0.1529807 , -0.11795644,  0.0101596 , -0.05307419, -0.15964694,
       -0.05431768,  0.01793781, -0.18792124, -0.13062844,  0.06617995,
       -0.13204704, -0.02890955,  0.18061101, -0.09911279, -0.08864116,
       -0.05147275, -0.11063503, -0.12750539, -0.06813971,  0.00510955,
       -0.01756637,  0.07310836,  0.01116222, -0.16218664, -0.05341996],
      dtype=float32)
```

Gambar 5. Hasil Vektor dari Kata bagus

3.2. Pengujian

1. Data pengujian
   Data yang digunakan pada penelitian ini 10.000 data berupa 5.000 data ulasan berlabel positif dan 5.000 data ulasan berlabel negatif. Proses pembentukan model memerlukan data yang seimbang karena apabila data tidak seimbang akan menjadikan model yang terbentuk hanya sensitif pada kelas yang dominan. Gambaran dari pembagian data dijelaskan pada Gambar 6.

Gambar 6. Pembagian Data Pemodelan

2. Skenario
   Skenario 1 merupakan pengujian dan analisa pengaruh kombinasi parameter *Word2Vec* terhadap nilai akurasi, sedangkan skenario 2 merupakan pengujian dan analisa pengaruh kombinasi parameter *Bi-LSTM* terhadap nilai akurasi. Gambaran umum skenario pengujian dijelaskan pada Gambar 7.

Gambar 7. Skenario Pengujian

Terdapat tujuh parameter dalam penelitan ini yang digunakan untuk pengujian. Ketujuh parameter tersebut dibandingkan kinerjanya terhadap model *Bidirectional Long Short Term Memory* (*Bi-LSTM*). Kelima paramter tersebut terbai menjadi 3 paramter untuk pengujian *Word2Vec* dan 2 parame ter untuk pengujian *Bi-LSTM*. Parameter model *Word2Vec* yang akan diujikan pada penelitian ini yaitu arsitektur model CBOW dan model Skipgram, metode evaluasi menggunakan *Hierarchical Softmax* dan *Negative Sampling*, sedangkan untuk dimensi menggunakan 100, 200, dan 300. Hasil kombinasi parameter model *Word2Vec* selanjutnya diproses menggunakan model *Bi-LSTM*.

Pada model *Bi-LSTM* terdapat 2 parameter berupa *dropout*, dan *learning rate*. Parameter *dropout* yang akan diuji adalah 0,2; 0,5; 0,7. Parameter *learning rate* menggunakan nilai 0,001; 0,0001; 0,00001. Keseluruhan parameter pada model *Word2Vec* dan parameter pada model *Bi-LSTM* dikombinasikan untuk mendapatkan model terbaik. Kinerja kombinasi parameter ditentukan dengan perhitungan

nilai akurasi. Nilai akurasi keseluruhan kombinasi yang didapatkan dibandingkan dan dianalisa untuk menentukan model terbaik. Berikut merupakan penjelasan dari masing-masing skenario:

1. Skenario 1
   Skenario 1 berisi pengujian dan analisa terhadap pengaruh kombinasi dari model *Word2Vec*. Berikut penjelasan masing-masing parameter dari skenario 1:
   a Pengujian dilakukan untuk mengetahui pengaruh arsitektur *Word2Vec* dalam mendapatkan nilai akurasi terbaik. Arsitektur *Word2Vec* yang digunakan pada proes pengujian yaitu model CBOW dan model *Skip-gram*. Eksperimen yang terbentuk pada penelitian ini sebanyak 108, teridiri dari model CBOW sebanyak 54 dan model *Skip-gram* sebanyak 54. Data masukan pada skenario ini berupa data hasil prapengolahan teks.
   b Pengujian dilakukan untuk mengetahui pengaruh metode evaluasi *Word2Vec* yang digunakan dalam mendapatkan nilai akurasi terbaik. Metode evaluasi *Word2Vec* yang digunakan pada proses pengujian antara lain, *Hierarchical Softmax* dan *Negative Sampling*. Eksperimen yang terbentuk pada penelitian ini sebanyak 108, terdiri dari model *Hierarchical Softmax* sebanyak 54 dan model *Negative Sampling* sebanyak 54. Data masukan pada skenario ini berupa data hasil prapengolahan teks.
   c Pengujian dilakukan untuk mengetahui pengaruh dimensi *Word2Vec* yang digunakan dalam mendapatkan nilai akurasi terbaik. Dimensi *Word2Vec* yang digunakan pada pada proses pengujian sebesar 100, 200, 300. Eksperimen yang terbentuk pada penelitian ini sebanyak 108, terdiri dari dimensi 100 sebanyak 36, 200 sebanyak 36, dan 300 sebanyak 36. Data masukan pada skenario ini berupa data hasil prapengolahan teks.
2. Skenario 2
   Skenario 2 berisi pengujian dan analisa terhadap pengaruh kombinasi dari model *Bi-LSTM*. Berikut penjelasan masing-masing parameter dari skenario 2:
   a Pengaruh nilai *dropout* yang digunakan untuk mendapatkan nilai akurasi terbaik. Nilai *dropout* yang digunakan pada proses pengujian sebesar 0,2; 0,3; 0,7. Eksperimen yang terbentuk pada penelitian ini sebanyak 108, terdiri dari dropout 0,2 sebanyak 36, dropout 0,5 sebanyak 36, dan *dropout* 0,7 sebanyak 36. Data masukan pada skenario berupa *array* hasil *Word2Vec* yang ada pada skenario 1.
   b Pengaruh nilai *learning rate* yang digunakan untuk mendapatkan nilai akurasi terbaik. Nilai *learning rate* yang digunakan pada proses pengujian sebesar 0,001; 0,0001; 0,0001. Eksperimen yang terbentuk pada penelitian ini sebanyak 108, terdiri dari *learning rate* 0,001 sebanyak 36, learning rate 0,0001 sebanyak 36, dan learning rate 0,00001 sebanyak 36. Data masukan pada skenario ini berupa *array* hasil *Word2Vec* yang ada pada skenario 1.

3.3. Analisa

Analisa nilai akurasi dari skenario pengujian sebanyak 108 kombinasi. Model terbaik dipilih dengan menganalisa nilai akurasi untuk kombinasi *Word2Vec* berupa arsitektur, metode evaluasi, dan dimensi serta BiLSTM berupa *dropout*, dan *learning rate*. Model terbaik yang terpilih selanjutnya digunakan untuk analisis sentimen objek wisata. Adapun Hasil pengujian dari skenario 1 dan 2 adalah sebagai berikut:

1. CBOW dan *Skipgram* merupakan arsitektur *Word2Vec* yang diujikan dan hasil rata-rata dihitung untuk setiap arsitektur *Word2Vec*, Dimana arsitektur *Word2Vec* CBOW menghasilkan rata-rata akurasi yang lebih baik dibandingkan dengan aristektur *Word2Vec Skipgram*.
2. *Hierarchical Softmax* dan *Negative Sampling* merupakan metode evaluasi *Word2Vec* yang diujikan. Hasil rata-rata dihitung untuk setiap metode evaluasi *Word2Vec* dengan menggunaka metode evaluasi *Word2Vec Hierarchical Softmax* menghasilkan rata-rata akurasi lebih baik dibandingkan dengan metode evaluasi *Word2Vec Negative Sampling*.
3. Dimensi 100, 200, 300 merupakan dimensi *Word2Vec* yang diujikan. Hasil rata-rata dihitung untuk setiap dimensi *Word2Vec*. Hasilnya bahwa dimensi *Word2Vec* 200 menghasilkan rata-rata akurasi yang lebih baik dibandingkan dengan dimensi 100 dan 300.
4. Nilai *dropout* yang diujikan antara lain 0,2; 0,5; 0,7. Hasil rata-rata dihitung untuk setiap jumlah *dropout*. *Dropout* dengan jumlah 0,5 menghasilkan rata-rata nilai akurasi yang lebih baik dari *dropout* dengan jumlah 0,2 dan 0,7.
5. Nilai 0,001; 0,0001; 0,00001 merupakan nilai *learning rate* yang diujikan. Hasil rata-rata dihitung untuk setiap nilai *learning rate* bahwa *learning rate* dengan nilai 0,0001 menghasilkan rata-rata akurasi lebih baik dibandingkan *learning rate* dengan nilai 0,001 dan 0,00001.

Nilai terbaik dari setiap skenario menghasilkan kombinasi parameter yaitu kombinasi model *Word2Vec* terdiri dari model CBOW, metode evaluasi *Hierarchical Softmax*, dan dimensi 200, serta *BiLSTM* yang terdiri dari dropout dengan nilai 0,5 dan *learning rate* dengan nilai 0,0001. Kombinasi tersebut menghasilkan nilai *akurasi* tertinggi dari keseluruhan 108 kombinasi yang hasilnya dengan nilai akurasi sebesar 96,86%, *precission* sebesar 96,53%, *Recall* sebesar 96,31%, F1 Measure sebesar 96,41%.

Penelitian ini menggunakan nilai akurasi sebagai acuan dalam menentukan model terbaik. Akurasi baik digunakan sebagai acuan penilaian kinerja algoritma jika dataset memiliki jumlah data *False* Negatif dan *Flase* Positif yang sangat mendekati. Karena itu nilai akurasi dijadikan acuan dalam membandingkan hasil penelitian ini dengan penelitian lainnya.

Hasil akurasi sebesar 96,86% itu lebih baik dibandingkan dengan penelitian [14] yang menggunakan metode pretraining kombinasi TF-IDF dan Seninfo dan menggunakan *deep learning Bi-LSTM*. Penelitian ini melakukan analisis sentiment data komentar berbahasa mandarin yang terdapat pada *e-commerce*. Penelitian lainnya [13] juga dengan yang mehasilkan kinerja tidak lebih baik dari penelitian ini dengan akurasi sebersar 91,08%. Peneliaian tersebut [13] menggunakan metode *GloveText* dan *Bi-LSTM* untuk analisis sentiment dari emosi lirik lagu.

4. KESIMPULAN

Kesimpulan dari hasil penelitian yang dilakukan mengenai analisis sentimen menggunakan model *Bi-LSTM* dan *Word2Vec* terhadap ulasan objek wisata Pulau Bali di tripadvisor.com berbahasa Indonesia adalah memperoleh model terbaik dari metode *Bi-LSTM* dan *Word2Vec* dengan akurasi sebesar 96,86%, precission sebesar 96,53%, Recall sebesar 96,31%, F1 Measure sebesar 96,41%. Adapun parameter *Word2Vec* dalam kontribusi yang menghasilkan model terbaik yaitu, arsitektur CBOW, metode evaluasi *hierarchical softmax*, dan dimensi 200. Sedangkan parameter *Bi-LSTM* yang berkontribusi menghasilkan model terbaik yaitu *dropout* 0,5 dan *learning rate* 0,0001.

Saran untuk penelitian selanjutnya adalah dapat menggunakan metode *pretraining* dan metode *deep learning* yang lain untuk membandingkan hasilnya dengan penelitian ini. Selain itu, penelitian selanjutnya juga dapat menggunakan metode transfer *learning* pada dataset lain yang serupa seperti ulasan hotel, ulasan restoran, dan lainnya. Transfer *learning* dilakukan dengan memanfaatkan model terbaik penelitian ini terhadap dataset yang lain yang sejenis dengan menggunakannya sebagai *starting point*, memodifikasi dan mengubah parameter sesuai dataset yang baru.

REFERENSI

[1] M. Antara and M. S. Sumarniasih, "Role of Tourism in Economy of Bali and Indonesia," *Journal of Tourism and Hospitality Management*, vol. 5, no. 2, pp. 34–44, 2017.

[2] X. Liu, F. Mehraliyev, C. Liu, and M. Schuckert, "The Roles of Social Media in Tourists' Choices of Travel Components," *Tourist Studies*, vol. 20, no. 1, pp. 27–48, 2020.

[3] Z. Ke, J. Sheng, Z. Li, W. Silamu, and Q. Guo, "Knowledge-Guided Sentiment Analysis Via Learning From Natural Language Explanations," *IEEE Access*, vol. 9, pp. 3570–3578, 2021.

[4] T. Iqbal and S. Qureshi, "The Survey: Text Generation Models in Deep Learning," *Journal of King Saud University - Computer and Information Sciences*, pp. 1–14, apr 2020.

[5] A. Yenter, "Deep CNN-LSTM with Combined Kernels from Multiple Branches for IMDb Review Sentiment Analysis," in *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, 2017, pp. 540–546.

[6] H. Ghulam, F. Zeng, W. Li, and Y. Xiao, "Deep Learning-Based Sentiment Analysis for Roman Urdu Text," *Procedia Computer Science*, vol. 147, pp. 131–135, 2019.

[7] D. Li and J. Qian, "Text Sentiment Analysis Based on Long Short-Term Memory," in *2016 First IEEE International Conference on Computer Communication and the Internet (ICCCI)*. IEEE, oct 2016, pp. 471–475.

[8] D. I. Af'idah, R. Kusumaningrum, and B. Surarso, "Long Short Term Memory Convolutional Neural Network for Indonesian Sentiment Analysis Towards Touristic Destination Reviews," in *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*. IEEE, sep 2020, pp. 630–637.

[9] N. Chen and P. Wang, "Advanced Combined LSTM-CNN Model for Twitter Sentiment Analysis," in *2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS)*. IEEE, nov 2018, pp. 684–687.

[10] Jiddy Abdillah, Ibnu Asror, and Yanuar Firdaus Arie Wibowo, "Emotion Classification of Song Lyrics Using Bidirectional LSTM Method with GloVe Word Representation Weighting," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 4, pp. 723–729, aug 2020.

[11] A. Aziz Sharfuddin, M. Nafis Tihami, and M. Saiful Islam, "A Deep Recurrent Neural Network with BiLSTM Model for Sentiment Classification," in *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*. IEEE, sep 2018, pp. 1–4.

[12] K. Zhang, W. Song, L. Liu, X. Zhao, and C. Du, "Bidirectional Long Short-Term Memory for Sentiment Analysis of Chinese Product Reviews," in *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*. IEEE, jul 2019, pp. 1–4.

[13] W. Li, F. Qi, M. Tang, and Z. Yu, "Bidirectional LSTM with Self-Attention Mechanism and Multi-Channel Features for Sentiment Classification," *Neurocomputing*, vol. 387, pp. 63–77, apr 2020.

[14] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment Analysis of Comment Texts Based on BiLSTM," *IEEE Access*, vol. 7, pp. 51 522–51 532, 2019.

[15] R. P. Nawangsari, R. Kusumaningrum, and A. Wibowo, "Word2Vec for Indonesian Sentiment Analysis Towards Hotel Reviews: An Evaluation Study," *Procedia Computer Science*, vol. 157, pp. 360–366, 2019.

[16] D. Jatnika, M. A. Bijaksana, and A. A. Suryani, "Word2Vec Model Analysis for Semantic Similarities in English Words," in *Procedia Computer Science*, vol. 157. Elsevier B.V., 2019, pp. 160–167.

[17] C. Zhang, X. Wang, S. Yu, and Y. Wang, "Research on Keyword Extraction of Word2vec Model in Chinese Corpus," in *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*. IEEE, jun 2018, pp. 339–343.

[18] I. Guyon, *A Scaling Law for The Validation-Set Training-Set Size Ratio*, 1997.