

BIDIRECTIONAL LONG SHORT TERM MEMORY DAN WORD2VEC
UNTUK ANALISIS SENTIMEN TERHADAP
ULASAN DESTINASI WISATA PULAU BALI



LAPORAN PENELITIAN

Sebagai Bukti Pengamalan Tri Dharma Perguruan Tinggi

Oleh:

- | | |
|---|-------------------|
| 1. Dwi Intan Af'idah, S.T., M.Kom. | 11.020.470 |
| 2. Dairoh, M.Sc. | 04.014.178 |
| 3. Susi Nurindah Sari | 19090102 |

PROGRAM STUDI SARJANA TERAPAN TEKNIK INFORMATIKA

POLITEKNIK HARAPAN BERSAMA

AGUSTUS 2021

SK Penelitian: SK Pengabdian: 090.05/PHB/V/2021, Tanggal: 31 Mei 2021
Nomor kontrak: 040716/P3M.PHB/VI/2021

BIDIRECTIONAL LONG SHORT TERM MEMORY DAN WORD2VEC
UNTUK ANALISIS SENTIMEN TERHADAP
ULASAN DESTINASI WISATA PULAU BALI



LAPORAN PENELITIAN

Sebagai Bukti Pengamalan Tri Dharma Perguruan Tinggi

Oleh:

- | | |
|---|-------------------|
| 1. Dwi Intan Af'idah, S.T., M.Kom. | 11.020.470 |
| 2. Dairoh, M.Sc. | 04.014.178 |
| 3. Susi Nurindah Sari | 19090102 |

PROGRAM STUDI SARJANA TERAPAN TEKNIK INFORMATIKA
POLITEKNIK HARAPAN BERSAMA
AGUSTUS 2021

SK Penelitian: 090.05/PHB/V/2021, Tanggal: 31 Mei 2021
Nomor kontrak: 047.16/P3M.PHB/V/2021

**HALAMAN PERSETUJUAN
LAPORAN PENELITIAN**

***BIDIRECTIONAL LONG SHORT TERM MEMORY DAN WORD2VEC
UNTUK ANALISIS SENTIMEN TERHADAP ULASAN DESTINASI
WISATA PULAU BALI***

Sebagai Bukti Pengamalan Tri Dharma Perguruan Tinggi

Oleh:

Nama	NIPY
1. Dwi Intan Af'idah, S.T., M.Kom.	11.020.470
2. Dairoh, M.Sc.	04.014.178
3. Susi Nurindah Sari	19090102

Tegal, Agustus 2021

Menyetujui,

Ketua Program Studi DIV Teknik Informatika
POLITEKNIK HARAPAN BERSAMA



Slamet Wiyono, S.Pd, M.Eng.
NIPY. 08.015.222

Ketua P3M
POLITEKNIK HARAPAN BERSAMA



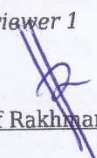
Kusnadi, M.Pd.
NIPY. 014.015.217

**HALAMAN PENGESAHAN
LAPORAN PENELITIAN**

- 1. Judul** : BIDIRECTIONAL LONG SHORT TERM MEMORY DAN WORD2VEC
UNTUK ANALISIS SENTIMEN TERHADAP ULASAN DESTINASI
WISATA PULAU BALI
- 2. Ketua Peneliti**
- a. Nama Lengkap : Dwi Intan Afidah, ST., M. Kom
 - b. NIDN : 0620089203
 - c. NIPY : 11.020.470
 - d. Jabatan Fungsional : -
 - e. Program Studi : Sarjana Terapan Teknik Informatika
 - f. Alamat e-mail :
- 3. Jumlah Anggota** : 3
- Nama Anggota 1 : Dairoh, M.Sc
 - Nama Mahasiswa 1 : Susi Nurindah sari
 - Nama Mahasiswa 2 : Susi Nurindah sari
- Biaya Penelitian** : Rp. 3,457,000

Tegal, Agustus 2021

Reviewer 1


Arif Rakhman, SE, S.Pd, M.Kom

NIPY. 05.016.291

Menyetujui,

Ketua Prodi Sarjana Terapan Teknik
Informatika

Politeknik Harapan Bersama


Spandhi Wicaksono, S. Pd.M. Eng

NIPY. 08.015.222

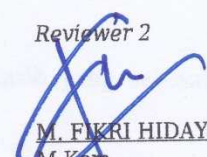
Mengetahui,

Wakil Direktur 1
Politeknik Harapan Bersama


Apt. Heru Nurcahyo, S.Farm., M.Sc

NIPY. 10.007.038

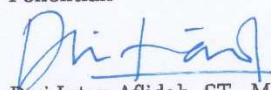
Reviewer 2


M. FIKRI HIDAYATTULLAH,
M. Kom

NIPY. 09.016.307

Ketua Tim Pelaksana

Penelitian


Dwi Intan Afidah, ST., M. Kom

NIPY. 11.020.470

Mengesahkan,

Ketua P3M
Politeknik Harapan Bersama


Kushadi, M.Pd

NIPY. 04.015.217

PRAKATA

Dengan memanjatkan puji syukur kehadirat Allah SWT, Tuhan Yang Maha Pengasih dan Maha Penyayang yang telah melimpahkan segala rahmat, hidayah dan inayah-Nya hingga terselesaikannya laporan penelitian yang dengan judul “*Bidirectional Long Short Term Memory* untuk Analisis Sentimen terhadap Ulasan Destinasi Wisata Pulau Bali”.

Laporan Penelitian ini merupakan dibuat sebagai bentuk tanggung jawab atas kegiatan penelitian yang di biaya oleh institusi sebagai salah satu kewajiban yang harus dilaksanakan oleh setiap dosen untuk memenuhi salah satu dharma pada Tri Dharma Perguruan Tinggi, yakni pada Politeknik Harapan Bersama. Selama Pelaksanaan dan penyusunan laporan penelitian tersusun, peneliti menyadari banyak pihak yang telah memberikan bantuan dan dukungan.

Pada kesempatan ini, peneliti mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Nizar Suhendra, S.E.MPP selaku Direktur Politeknik Harapan Bersama.
2. Bapak Kusnadi, M.Pd Selaku Ketua P3M Politeknik Harapan Bersama Tegal.
3. Bapak Slamet Wiyono, S.Pd, M.Eng selaku Ketua Program Studi DIV Teknik Informatika yang telah memberikan dukungan sarana maupun prasarana untuk penyusunan proposal penelitian.

Semoga laporan Penelitian ini dapat diterima dan memberikan sumbangan untuk pengembangan ilmu pengetahuan dan teknologi.

Tegal, Agustus 2021

Peneliti

PERNYATAAN

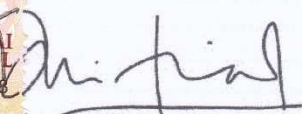
Dengan ini kami menyatakan bahwa:

1. Penelitian ini tidak pernah dibuat oleh peneliti lain dengan tema, judul, isi, metode, objek penelitian yang sama.
2. Penelitian ini bukan merupakan karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi.
3. Dalam penelitian ini juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Tegal, Agustus 2021



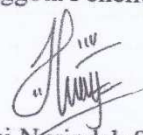
Ketua Tim Peneliti


Dwi Intan Afidah, S.T., M.Kom.
NIPY. 11.020.470

Anggota Peneliti 1


Dairoh, M.Sc.
NIPY. 04.14.178

Anggota Peneliti 2


Susi Nurindah Sari
NIPY. 19090102

DAFTAR ISI

	Halaman
HALAMAN SAMPUL.....	i
HALAMAN PERSETUJUAN	Error! Bookmark not defined.
HALAMAN PENGESAHAN	Error! Bookmark not defined.
PRAKATA	iii
PERNYATAAN	Error! Bookmark not defined.
DAFTAR ISI	vi
ABSTRAK.....	vii
BAB I.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah	3
1.3 Pembatasan Masalah.....	3
1.4 Tujuan	4
1.5 Manfaat	4
BAB II	5
2.1 Tinjauan Pustaka.....	5
2.2 Dasar Teori	7
2.2.1 Analisis Sentimen.....	7
2.2.2 Prapengolahan Teks	9
2.2.3 Word2Vec	10
2.2.4 Bidirectional Long Short Term Memory	12
BAB III	6
3.1 Bahan Penelitian	6
3.2 Alat Penelitian.....	6
3.3 Prosedur Penelitian	7
3.3.1 Pengumpulan Data.....	7
3.3.2 Perancangan Alur Program	7
3.3.2 <i>Preprocessing</i>	9
3.3.3 <i>Pretraining</i> Metode <i>Word2Vec</i>	10
3.3.4 <i>Pembagian Data</i>	11
3.3.5 Pembentukan Model	11
3.3.6 Evaluasi.....	12
BAB IV	14
4.1 Hasil Penelitian	14
4.2 Pembahasan	14
4.4.1 Skenario Penelitian	14
2.4.2 Hasil dan Analisa Penelitian	17
4.3 Luaran yang Dicapai.....	23
BAB IV	24
5.1 Kesimpulan	24
5.2 Saran	25
DAFTAR PUSTAKA.....	26
Lampiran 1.....	28
Lampiran 2.....	31
Lampiran 3.....	33

ABSTRAK

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. Hal tersebut mengakibatkan perlu dilakukannya identifikasi objek wisata Pulau Bali yang diminati masyarakat dengan cara menganalisis ulasan masyarakat yang tersebar di internet mengenai suatu objek wisata. Penelitian ini bertujuan melakukan analisis sentimen terhadap ulasan objek wisata di Pulau Bali menggunakan metode Bi-LSTM (*Bidirectional Long Short Term Memory*) dan *Word2Vec*, sehingga diperoleh model terbaik dengan mengetahui kinerja dari kombinasi parameter Bi-LSTM dan *Word2Vec*. Model *Word2Vec* yang menghasilkan akurasi paling baik untuk melakukan klasifikasi dengan menggunakan parameter arsitektur CBOW, metode evaluasi *hierarchical softmax*, dan dimensi 200. Model BiLSTM menghasilkan akurasi terbaik untuk melakukan klasifikasi dengan menggunakan parameter *dropout 0,5*, dan *learning rate 0,0001*. Kombinasi parameter model BiLSTM dan model *Word2Vec* yang terbaik menghasilkan nilai akurasi sebesar 96,86% dan digunakan untuk membuat klasifikasi analisis sentimen objek wisata berbahasa Indonesia.

Kata Kunci: Analisis Sentimen, Bidirectional Long Short Term Memory, Word2Vec

BAB I

PENDAHULUAN

1.1 Latar Belakang

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.

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].

LSTM selanjutnya dikembangkan menjadi *Bidirectional Long Short Term Memory* (Bi-LSTM) 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 [7] [8] [9]. Masalah lain yang muncul dalam analisis sentimen adalah penentuan metode *pretraining* yang tepat agar diperoleh model yang 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 [10].

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 pada ulasan objek wisata di Pulau Bali dibutuhkan bagi pemerintah agar objek wisata yang masih kurang diminati memperoleh perhatian lebih untuk dikembangkan. Sedangkan bagi wisatawan, analisis sentimen dapat menjadi referensi dalam mempertimbangan kunjungan wisata ke Pulau Bali.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka permasalahan dalam penelitian ini adalah:

1. Kementerian Pariwisata mengalami kesulitan dalam manajemen pengembangan pariwisata di Pulau Bali dikarenakan jumlah objek wisata di Pulau Bali yang banyak membuat pemerintah kesulitan dalam menentukan skala prioritas objek wisata yang perlu dikembangkan.
2. Wisatawan masih kesulitan dalam mencari informasi terkait tempat wisata baru di Pulau Bali yang sesuai dengan keinginan dan direkomendasikan banyak orang.
3. Belum adanya penelitian yang menggunakan metode Bi-LSTM dan *Word2Vec* untuk analisis sentimen teks berbahasa Indonesia terhadap ulasan objek wisata di Pulau Bali.
4. Diperlukan model *deep learning* terbaik dengan tingkat akurasi terbaik untuk penilaian sentimen teks berbahasa Indonesia terhadap ulasan objek wisata di Pulau Bali.

1.3 Pembatasan Masalah

Agar lingkup permasalahan tidak melebar, maka usulan penelitian ini dibatasi pada beberapa poin berikut:

1. Hasil usulan penelitian berupa model terbaik dari metode Bi-LSTM dan *Word2Vec* untuk penilaian sentimen teks berbahasa Indonesia terhadap ulasan objek wisata di Pulau Bali.
2. Data yang digunakan pada penelitian ini merupakan ulasan objek wisata Pulau Bali berbahasa Indonesia pada situs [tripadvisor.com](https://www.tripadvisor.com) yang diperoleh dari penelitian [11].

1.4 Tujuan

Tujuan dari usulan penelitian ini adalah:

1. melakukan analisis sentimen terhadap ulasan objek wisata di Pulau Bali menggunakan metode Bi-LSTM dan *Word2Vec*.
2. diperolehnya model terbaik dengan mengetahui kinerja arsitektur *Word2Vec*, metode evaluasi *Word2Vec*, dan dimensi *Word2Vec* serta kinerja *dropout* Bi-LSTM, *pooling* Bi-LSTM, dan Bi-LSTM *learning rate*.

1.5 Manfaat

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. mengetahui sentimen masyarakat terhadap suatu objek wisata Pulau Bali yang selanjutnya dapat menjadi referensi dalam melakukan pertimbangan untuk pengembangan pariwisata Pulau Bali bagi Kementerian Pariwisata dan pelaku industri wisata di Pulau Bali.
2. mengetahui sentimen masyarakat terhadap suatu objek wisata Pulau Bali yang selanjutnya dapat menjadi referensi dalam melakukan penentuan objek wisata yang akan dikunjungi bagi wisatawan.
3. menjadi referensi bagi peneliti yang akan melakukan penelitian dengan tema yang serupa.
4. terealisasinya tri dharma perguruan tinggi.

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Perkembangan *machine learning* klasik menjadi *deep learning* menjadikan teknik analisis sentimen juga mulai berkembang menggunakan metode *deep learning*. Beberapa penelitian analisis sentimen baik yang menggunakan *machine learning* klasik maupun *deep learning* dapat dilihat pada Tabel 2.1.

Tabel 2.1 Penelitian Analisis Sentimen Terdahulu

No.	Peneliti	Rekomendasi Metode	Keterangan
1.	(Ghulam dkk.. 2019)	<i>Long Short Term Memory (LSTM)</i>	Akurasi LSTM lebih baik dibandingkan SVM
2.	(Li dan Qian 2016)	<i>Long Short Term Memory (LSTM)</i>	LSTM memiliki kinerja yang lebih baik dibandingkan <i>Recurrent Neural network (RNN)</i>
3.	(Xu dkk., 2018)	<i>Bidirectional Long Short Term Memory (Bi-LSTM)</i>	Bi-LSTM memiliki kinerja lebih baik dibandingkan metode <i>deep learning</i> lainnya.
4.	(Abdillah dkk., 2020)	<i>Bidirectional Long Short Term Memory (Bi-LSTM)</i>	Bi-LSTM dan <i>Glove Word</i> memiliki kinerja lebih baik dibandingkan metode <i>deep learning</i> lainnya.
5.	(Af'idah dkk., 2020)	<i>Long Short Term Memory - Convolutional Neural network (LSTM-CNN)</i>	LSTM-CNN dan <i>Word2Vec</i> terhadap analisis sentimen ulasan objek wisata Pulau Bali

Penelitian [6] membandingkan metode *machine learning* klasik dan *deep learning* dilakukan untuk analisis sentimen pada ulasan *e-marketing* berupa teks Roman Urdu. *Long Short Term Memory (LSTM)* merupakan metode dari *deep*

learning yang dibandingkan dengan tiga metode *machine learning* klasik yaitu, *Naïve Bayes* (NB), *Random Forest* (RF), dan *Support Vector Machine* (SVM). Hasil penelitian ini menunjukkan bahwa LSTM memiliki akurasi paling baik sebesar 95% dibandingkan SVM sebesar 92%, RF sebesar 88%, dan NB sebesar 77%. LSTM juga dikatakan sebagai metode yang efisien untuk analisis sentimen karena tidak membutuhkan seleksi fitur [6].

Long Short Term Memory (LSTM) merupakan salah satu pengembangan *Recurrent Neural network* (RNN) untuk mengatasi masalah difusi gradien. Penelitian [12] 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. Jumlah data JD.COM sebanyak 4.000 data latih dan 1.800 data uji, sedangkan jumlah data Ctrip Travel dari Cina sebanyak 6.000 data latih dan 2.000 data uji, serta jumlah data ulasan film sebanyak 12.500 data latih dan 12.500 data uji. 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 [12].

Penelitian [9] membuktikan bahwa Bi-LSTM memiliki kinerja lebih baik dibandingkan metode *deep learning* lainnya dalam melakukan analisis sentimen berbahasa mandarin. Penelitian ini melakukan penilaian sentimen terhadap komentar yang terdapat pada website *e-commerce* dan website *travelling*. *Pretraining* yang digunakan pada penelitian ini adalah kombinasi dari TF-IDF dan Seninfo. Hasil penelitian ini menunjukkan metode Bi-LSTM dengan *Pretraining* TF-IDF dan Seninfo memiliki akurasi sebesar 91,54% yang merupakan akurasi tertinggi diantara metode *deep learning* lainnya, yaitu sebesar 91,54% [9].

Penelitian [13] melakukan klasifikasi emosi pada teks berbahasa Indonesia menggunakan metode Bi-LSTM dan *pretraining Glove Word*. *Dataset* yang digunakan untuk melakukan klasifikasi emosi merupakan data yang 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% [13].

Kombinasi dari dua metode *deep learning* atau lebih dapat dilakukan pada suatu analisis sentimen. Penelitian [11] menggunakan dua metode *deep learning* sekaligus yakni 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 pada penelitian ini menunjukkan adanya perbaikan akurasi dari metode LSTM-CNN dibandingkan metode LSTM tunggal [11].

Pada penelitian sebelumnya, Bi-LSTM menunjukkan performa yang lebih baik daripada LSTM biasa. 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. Memperhatikan keunggulan dari Bi-LSTM dan *Word2Vec*, maka metode Bi-LSTM digunakan untuk analisis sentimen teks berbahasa Indonesia dengan metode *pretraining* menggunakan *Word2Vec* dimana teks pada penelitian ini berasal dari ulasan objek wisata Pulau Bali di situs [tripadvisor.com](https://www.tripadvisor.com).

2.2 Dasar Teori

2.2.1 Analisis Sentimen

Natural Language Processing (NLP) mempelajari dan mengembangkan algoritma dan sistem yang memungkinkan komputer memahami dan melakukan tugas yang melibatkan bahasa manusia. NLP disebut juga linguistik komputasional, *computer speech* dan pengolahan bahasa. NLP dapat menganalisis bahasa manusia baik dalam bentuk tertulis maupun suara sehingga didapatkan informasi yang berguna. NLP bertujuan mengatasi masalah komputer dalam memahami bahasa alami manusia yang memiliki peraturan gramatikal dan semantik dengan cara mengubah bahasa alami manusia menjadi representasi data yang dapat dipahami dan dapat diolah oleh komputer (Pustejovsky dan Stubbs, 2012).

Analisis sentimen dikenal pula dengan *opinion mining* yang bertugas menganalisis opini, ekspresi dan perasaan para audiens terhadap suatu entitas. Opini selanjutnya diteliti termasuk positif, negatif ataupun netral. Analisis sentimen mampu menganalisis data teks seperti *tweet*, ulasan dalam media sosial, blog, dan *website* (Pang dan Lee, 2008). Analisis sentimen berupa teks ulasan bertujuan untuk menentukan arah ekspresif ulasan pengguna atau pelanggan. Penelitian analisis sentimen menjadi populer dikarenakan meningkatnya kebutuhan untuk menganalisis informasi tersembunyi dari ulasan pelanggan atau pengguna produk yang tersebar di media *online* dan situs internet dalam bentuk data yang tidak terstruktur (Luo dkk., 2016).

Analisis sentimen melakukan analisis data menggunakan metode klasifikasi. Klasifikasi memproses data dengan mengelompokkan data kedalam suatu kelas yang telah didefinisikan sebelumnya. Tujuan utama analisis sentimen adalah menemukan polaritas dari suatu data, sehingga proses analisis data dengan bentuk klasifikasi digunakan untuk memprediksi suatu data masuk kedalam kelas positif atau negatif atau netral. Terdapat tiga level dalam analisis sentimen yang menjadi acuan dalam sebuah penelitian, yaitu level dokumen, level kalimat dan level aspek (Liu, 2015).

Level dokumen memiliki tujuan untuk mengklasifikasikan satu keseluruhan dokumen termasuk kelas positif atau kelas negatif. Contoh analisis sentimen opini publik terhadap suatu tokoh, maka hasilnya berupa sentimen positif atau negatif dari tokoh tersebut berdasarkan satu kesatuan komentar. Kelebihan level dokumen adalah mampu menentukan politas secara utuh (Liu, 2015).

Level kalimat menentukan kalimat yang menjadi kalimat sentimen terlebih dahulu. Kalimat sentimen selanjutnya dianalisis apakah kalimat termasuk kelas positif atau kelas negatif. Diasumsikan dalam suatu dokumen terdapat beberapa kalimat, kemudian ditentukan lebih dulu mana yang merupakan kalimat opini yang berarti bukan kalimat fakta, selanjutnya ditentukan kalimat opini tersebut termasuk ke dalam kelas positif atau negatif (Liu, 2015).

Level aspek merupakan level terendah dimana proses analisis dilakukan lebih mendalam. Pada level dokumen dan kalimat penentuan polaritas dokumen tidak mampu menggambarkan target opini. Dalam suatu dokumen ulasan bisa saja terdapat lebih dari satu entitas yang menjadi target opini dimana masing-masing target opini berdiri sendiri dalam penentuan polaritasnya. Hal ini memungkinkan dalam satu dokumen dapat masuk kelas positif dan kelas negatif sekaligus. Misalkan terdapat kalimat berikut “tempat ini sangat indah, tetapi jalan menuju ke sana cukup berbahaya” kalimat ini menunjukkan terdapat dua entitas dari ulasan, yaitu terkait tempat dan akses. Pada kalimat tersebut aspek tempat termasuk kelas positif, sedangkan aspek akses termasuk kelas negatif (Liu, 2015).

2.2.2 Prapengolahan Teks

Data teks merupakan data yang tidak beraturan karena terdapat perulangan kata dan munculnya banyak kata yang tidak berkontribusi pada analisis data. Prapengolahan teks perlu dilakukan untuk membersihkan data teks terlebih dahulu sebelum dilakukan proses analisis sentimen. Prapengolahan teks menghilangkan data yang tidak konsisten, data yang duplikat, dan data yang tidak berpengaruh terhadap polaritas suatu dokumen. Tahapan proses prapengolahan teks antara lain:

1. *Case folding*

Proses mengubah semua karakter huruf pada sebuah kalimat menjadi huruf kecil atau huruf besar disebut *case folding*. *Case folding* yang dilakukan pada penelitian ini yaitu mengubah seluruh *dataset* menjadi huruf kecil. Huruf kapital biasanya terdapat pada setiap awal kalimat seperti “Udara di tempat ini sangat sejuk”, menggunakan *case folding* kalimat tersebut berubah menjadi “udara di tempat ini sangat sejuk”. Tujuan utama *case folding* adalah agar kata “udara” tidak lagi mempunyai dua bentuk yaitu, “Udara”, dan “udara”, namun hanya memiliki satu bentuk huruf kecil saja (Hidayatullah dan Ma’arif, 2016).

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*. Contoh karakter ilegal yang dihilangkan antara lain %, &, >, (, {, }, 1-9, @uluwatu, <http://tripadvisor.com> (Symeonidis dkk., 2018).

3. Stopword Removal

Stopword Removal merupakan tahap pengambilan kata-kata penting dan membuang kata-kata yang dianggap tidak penting. Cara untuk membuang kata yang tidak penting disebut *stopword removal*. *Stopword removal* bertujuan untuk menghilangkan kata-kata yang sering muncul namun tidak memiliki kontribusi dalam proses analisis data. *Stopword removal* berusaha memperkecil dimensi data dan mempercepat waktu komputasi (Symeonidis dkk., 2018). Contoh kata yang tidak penting di bahasa Indonesia seperti kata “dan”, ”yang”, ”di”, ”ke”.

4. Padding

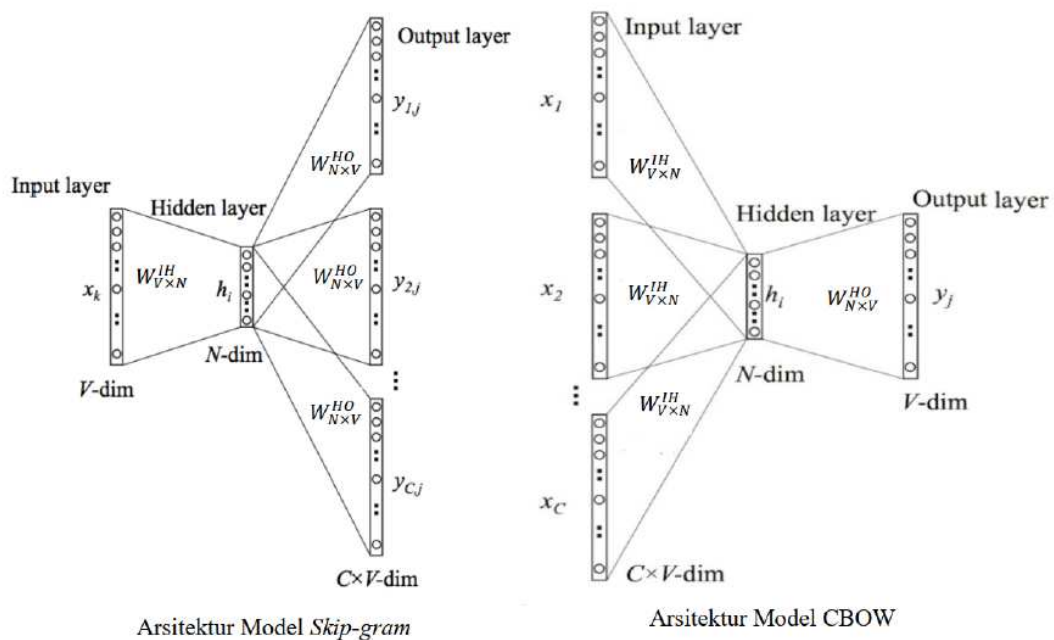
Proses pembelajaran yang dilakukan oleh *neural network* memerlukan masukan data dengan panjang yang sama. *Padding* merupakan proses yang dilakukan untuk membuat *input* mempunyai panjang yang sama dengan cara menambahkan kata “<pad>”. *Dataset* pada penelitian ini memiliki panjang teks yang berbeda-beda. Oleh karena itu, perlu dilakukan *padding* agar vektor memiliki panjang yang sama sebelum diproses pada *neural network* (Giménez dkk., 2020).

2.2.3 Word2Vec

Model *Word2Vec* merupakan model *word embedding* yang diusulkan oleh Mikolov, Sutskever, Chen, Corrado, dan Dean pada tahun 2013. *Word2Vec* memiliki kemampuan membaca dan mengolah teks dalam ukuran yang besar dan mengubah setiap kata menjadi vektor. Model *Word2Vec* mampu memahami sintaks

dan makna semantik kata dari bahasa alami kemudian merepresentasikan setiap kata dengan sebuah vektor. *Word2Vec* mencapai kinerja terbaik dalam NLP dengan mengelompokkan kata serupa yang memiliki vektor yang sama (Al-Amin dkk., 2017).

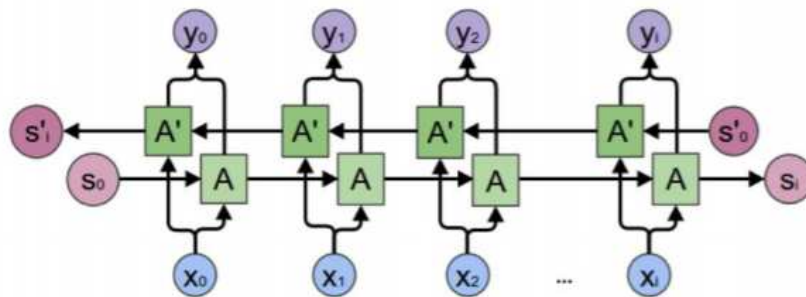
Neural Network digunakan pada model *Word2Vec* untuk menghasilkan *output* berupa ruang vektor dari *input* yang berupa korpus teks. Model *Word2Vec* memiliki dua jenis arsitektur bernama *Continuous Bag of Words* (CBOW) dan *Skip-gram*. Arsitektur CBOW memprediksi kata saat ini berdasarkan konteks, sedangkan arsitektur *Skip-gram* memprediksi dalam jangkauan sebelum atau setelah kata sekarang dimana kata sekarang merupakan *input*, sehingga arsitektur *Skip-gram* dianggap sebagai arsitektur yang efisien dalam mempelajari vektor kata yang tidak terstruktur dalam jumlah besar (Nawangsari dkk., 2019). Arsitektur *Skip-gram* dan CBOW dapat dilihat pada Gambar 2.1.



Gambar 0.1 *Skip-gram* dan CBOW (Nawangsari dkk., 2019)

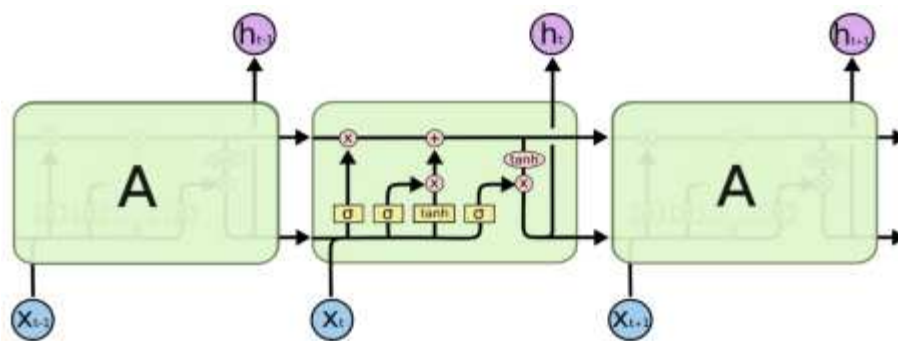
2.2.4 Bidirectional Long Short Term Memory

Bidirectional Long Short Term Memory (Bi-LSTM) merupakan salah satu jenis dari pengembangan *Long Short Term Memory* (LSTM). *Input* yang dimasukkan ke dalam BiLSTM ada 2 jenis yaitu *input forward* dan *input backward*. *Output* dari lapisan ini umumnya digabungkan menjadi satu. Dengan *layer* ini, model dapat mempelajari informasi masa lalu (*past*) dan informasi masa mendatang (*future*) untuk tiap sekuen *input*.

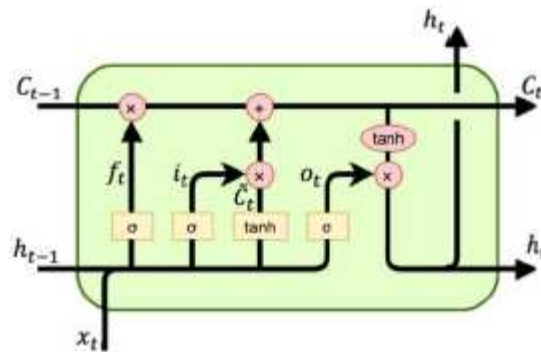


Gambar 0.3 Arsitektur Bi-LSTM (Olah, 2015)

Long Short Term Memory (LSTM) dikenalkan oleh Hochreiter dan Schmidhuber (1997) untuk mengatasi masalah difusi gradien *Recurrent Neural network* (RNN). LSTM merupakan salah satu variasi dari RNN yang dibuat untuk menghindari masalah mengingat informasi jangka panjang pada RNN. Jaringan perulangan RNN hanya menggunakan satu *layer* sederhana, yaitu *layer tanh* seperti, sedangkan LSTM memiliki empat *layer* pada perulangan modelnya seperti pada Gambar 2.3.

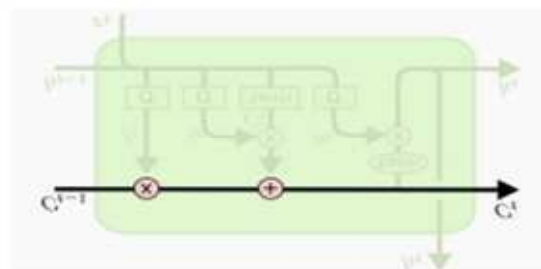


Gambar 0.3 Perulangan empat *layer* pada LSTM (Smagulova dan James, 2019)

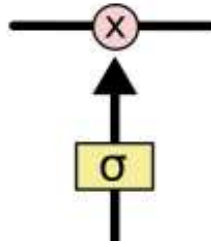


Gambar 0.4 Ilustrasi *Cell* LSTM (Smagulova dan James, 2019)

Cell pada LSTM dapat dilihat pada Gambar 2.4 yang terdiri dari *layer neuron* dilambangkan dengan persegi panjang, operasi *element-wise* dilambangkan dengan lingkaran. Panah hitam melambangkan aliran informasi di dalam *cell* dan antar *cell* maupun keluaran dari *cell* (*output h*). *Cell* LSTM mempunyai 2 hasil keluaran, yang pertama yaitu informasi yang sebenarnya *hidden state* (*ht*) yang diteruskan ke *cell* selanjutnya dan menjadi *input* dari *cell* selanjutnya, yang kedua yaitu *cell state* (*Ct*). *Cell state* merupakan kunci utama dari LSTM. *Cell state* merupakan garis horizontal yang menghubungkan semua *output layer* pada LSTM seperti terlihat pada Gambar 2.5. LSTM memiliki kemampuan untuk menambah dan menghapus informasi dari *cell state* yang disebut *gates*. *Gates* berfungsi mengatur informasi masuk secara opsional dengan menggunakan *sigmoid layer* yang digambarkan pada Gambar 2.6. Keluaran dari *sigmoid layer* menunjukkan informasi diteruskan atau diberhentikan. Angka 0 untuk diberhentikan dan angka 1 untuk diteruskan.



Gambar 0.5 *Cell state* pada LSTM (Smagulova dan James, 2019)



Gambar 0.6 Sigmoid *layer* pada LSTM (Smagulova dan James, 2019)

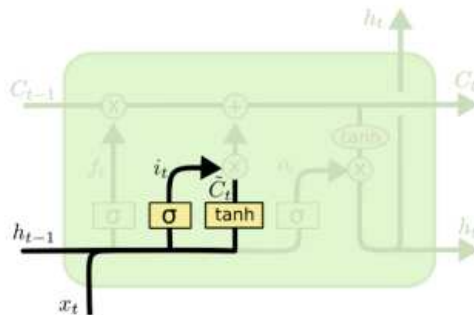
LSTM memiliki tiga jenis *gates* yaitu, *forget gate* (i_t), *input gate* (f_t), dan *output gate* (O_t). *Forget gate* berfungsi untuk memutuskan informasi yang dihapus dari *cell*. *Input gate* berfungsi untuk memutuskan nilai dari *input* untuk diperbaharui pada *state* memori. *Output gate* berfungsi untuk memutuskan apakah yang dihasilkan *output* sesuai dengan *input* dan memori pada *cell* atau tidak.

Proses berjalannya metode LSTM ada empat langkah (Smagulova dan James, 2019):

1. Langkah pertama bertujuan untuk memutuskan informasi yang akan disimpan di *cell state*. Langkah pertama memiliki dua bagian. Bagian pertama *sigmoid layer* yang bernama *input gate layer* untuk memutuskan nilai yang akan diperbarui. Bagian kedua *tanh layer* untuk membuat satu kandidat dengan nilai baru (\tilde{C}_t) yang dapat ditambahkan ke *cell state*. Tahap selanjutnya *output* dari *input gate layer* dan *tanh layer* akan digabungkan untuk memperbarui *cell state*. Langkah kedua digambarkan pada Gambar 2.7, sedangkan persamaan *input gate* diuraikan pada persamaan 2.1 dan persamaan kandidat baru diuraikan pada persamaan 2.2.

$$\tilde{C}_t = \tanh(W_c x(t) + U_c h(t-1)) + b_c \quad (2.1)$$

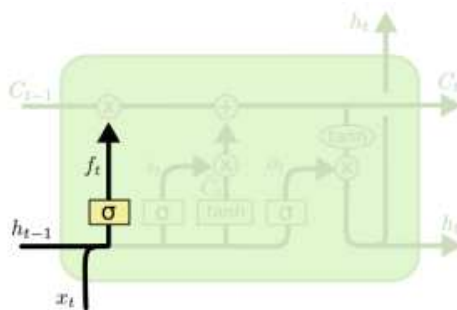
$$i_t = \sigma(W_i x(t) + U_i h(t-1)) + b_i \quad (2.2)$$



Gambar 0.7 *Input gate layer* dan *tanh layer* (Smagulova dan James, 2019)

- Langkah kedua LSTM memutuskan informasi apa yang akan dihapus dari *cell state*. Keputusan ini dibuat oleh *sigmoid layer* yang bernama *forget gate layer*. *Forget gate layer* akan memproses $h(t-1)$ dan $x(t)$ sebagai *input*, dan menghasilkan *output* berupa angka 0 atau 1 pada *cell state* C_{t-1} seperti pada Gambar 2.8, adapun persamaan *forget gate* diuraikan pada persamaan 2.3.

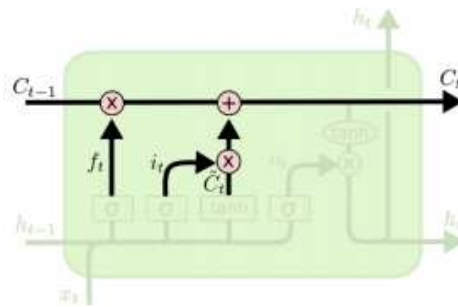
$$f_t = \sigma(W_f x(t) + U_f h(t-1) + b_f) \quad (2.3)$$



Gambar 0.82 *Forget gate layer* (Smagulova dan James, 2019)

- Langkah ketiga bertujuan untuk memperbarui *cell state* yang lama (C_{t-1}) menjadi *cell state* baru, C_t seperti pada Gambar 2.11. Perkalian *state* lama dengan f_t bertujuan untuk menghapus informasi yang sudah ditentukan sebelumnya pada langkah *forget gate layer*, selanjutnya ditambahkan dengan $i_t * \tilde{C}_t$ yang merupakan nilai baru dan digunakan untuk memperbarui *state*. Persamaan *cell state* diuraikan pada persamaan 2.5.

$$C_t = i_t * \tilde{C}_t + f_t * C_{t-1} \quad (2.5)$$

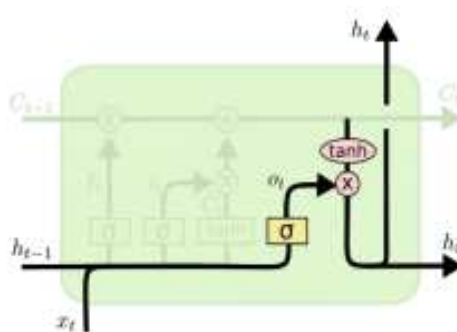


Gambar 0.93 Pembuatan *cell state* baru (Smagulova dan James, 2019)

4. Langkah keempat yang merupakan langkah terakhir dalam metode LSTM bertujuan untuk memutuskan hasil *output* seperti Gambar 2.10. *Output* harus sesuai dengan *cell state* yang telah diproses terlebih dahulu. Pertama, *sigmoid layer* memutuskan bagian dari *cell state* yang menjadi *output*. Kedua, *output* dari *cell state* dimasukkan ke dalam *tanh layer* (untuk mengganti nilai menjadi diantara -1 dan 1) dan dikalikan dengan *sigmoid gate* agar *output* yang dihasilkan sesuai dengan apa yang diputuskan sebelumnya. Persamaan *output gate* diuraikan pada Persamaan 2.6 dan 2.7. Seluruh keterangan notasi pada langkah LSTM dijelaskan pada Tabel 2.2.

$$O_t = \sigma(W_o x(t) + U_o h(t - 1) + b_o) \quad (2.6)$$

$$h_t = O_t * \tanh(C_t) \quad (2.7)$$



Gambar 0.10 Penentuan *output* (Smagulova dan James, 2019)

5. Pada Bi-LSTM hasil dari LSTM akan dioperasikan menggunakan operasi *sum* (penjumlahan) atau perkalian.

Tabel 0.1 Keterangan notasi dan definisi persamaan LSTM

Notasi	Definisi
t	<i>Time step $t = 1, 2, \dots$</i>
$x(t)$ atau x_t	Vektor masukan pada saat <i>time step</i> ke t
$h(t - 1)$	<i>Hidden state</i> pada <i>time step</i> ke $t-1$
h_t	<i>Hidden state time step</i> ke t
C_{t-1}	<i>Cell memory time step</i> ke $t-1$
C_t	<i>Cell memory time step</i> ke t
W_c	Bobot pada <i>candidate layer</i> untuk <i>input kata</i>
W_i	Bobot pada <i>input layer</i> untuk <i>input kata</i>
W_o	Bobot pada <i>output layer</i> untuk <i>input kata</i>
W_f	Bobot pada <i>forget layer</i> untuk <i>input kata</i>
U_c	Bobot pada <i>candidate layer</i> untuk <i>hidden state time step</i> sebelum
U_i	Bobot pada <i>input layer</i> untuk <i>hidden state time step</i> sebelum
U_o	Bobot pada <i>output layer</i> untuk <i>hidden state time step</i> sebelum
U_f	Bobot pada <i>forget layer</i> untuk <i>hidden state time step</i> sebelum
*	<i>Hadamard product</i>

BAB III METODE PENELITIAN

3.1 Bahan Penelitian

Bahan penelitian yang digunakan dalam penelitian ini berupa sekumpulan teks ulasan objek wisata Pulau Bali berbahasa Indonesia yang diambil dari situs [tripadvisor.com](https://www.tripadvisor.com) yang berasal dari penelitian sebelumnya yakni penelitian [11].

3.2 Alat Penelitian

Alat penelitian yang digunakan dalam penelitian ini meliputi perangkat keras dan perangkat lunak. Berikut spesifikasi dari alat digunakan:

1. Perangkat keras (*hardware*)

Perangkat keras yang dibutuhkan meliputi: komputer dengan spesifikasi minimum prosesor berfrekuensi 1 GHz, RAM 4 GB, dan *SSD 500GB*. *Printer* multifungsi minimal setara Canon MP140 dilengkapi dengan *scanner*.

2. Perangkat lunak (*software*)

Selain perangkat keras, dukungan perangkat lunak juga diperlukan dalam melakukan penelitian. Perangkat lunak yang digunakan meliputi:

Tabel 3.1 Tabel Kebutuhan Perangkat Lunak

Nama Perangkat Lunak	Fungsi
Python 3.7 x64, Anaconda 3	<i>Tools</i> untuk membuat program yang berbahasa Python
Google Collabs	<i>Tools</i> yang digunakan untuk melakukan proses pelatihan dan pengujian model
Scikit-Learn, Sastrawi dan Tensorflow 1.8	<i>library</i> yang digunakan pada proses klasifikasi sentimen menggunakan <i>machine learning</i>
Microsoft Office Word 2016	<i>Software</i> untuk membuat laporan penelitian.
Microsoft Office Excel 2016	<i>Software</i> untuk mengelola data yang akan diimpor ke dalam aplikasi hasil penelitian.
Microsoft Visio 2016	<i>Software</i> untuk merancang model, flowchart dan lain-lain
Mozilla Firefox, Chrome	<i>Browser</i>

3.3 Prosedur Penelitian

Prosedur yang dilakukan pada penelitian ini seperti ditunjukkan Gambar 3.1. Prosedur penelitian ini terdiri dari pengumpulan data, perancangan alur program, persiapan *dataset* terdiri dari *preprocessing* dan *pretraining Word2Vec*, pembagian data, pembentukan model, serta evaluasi.



Gambar 3.1 Prosedur Penelitian

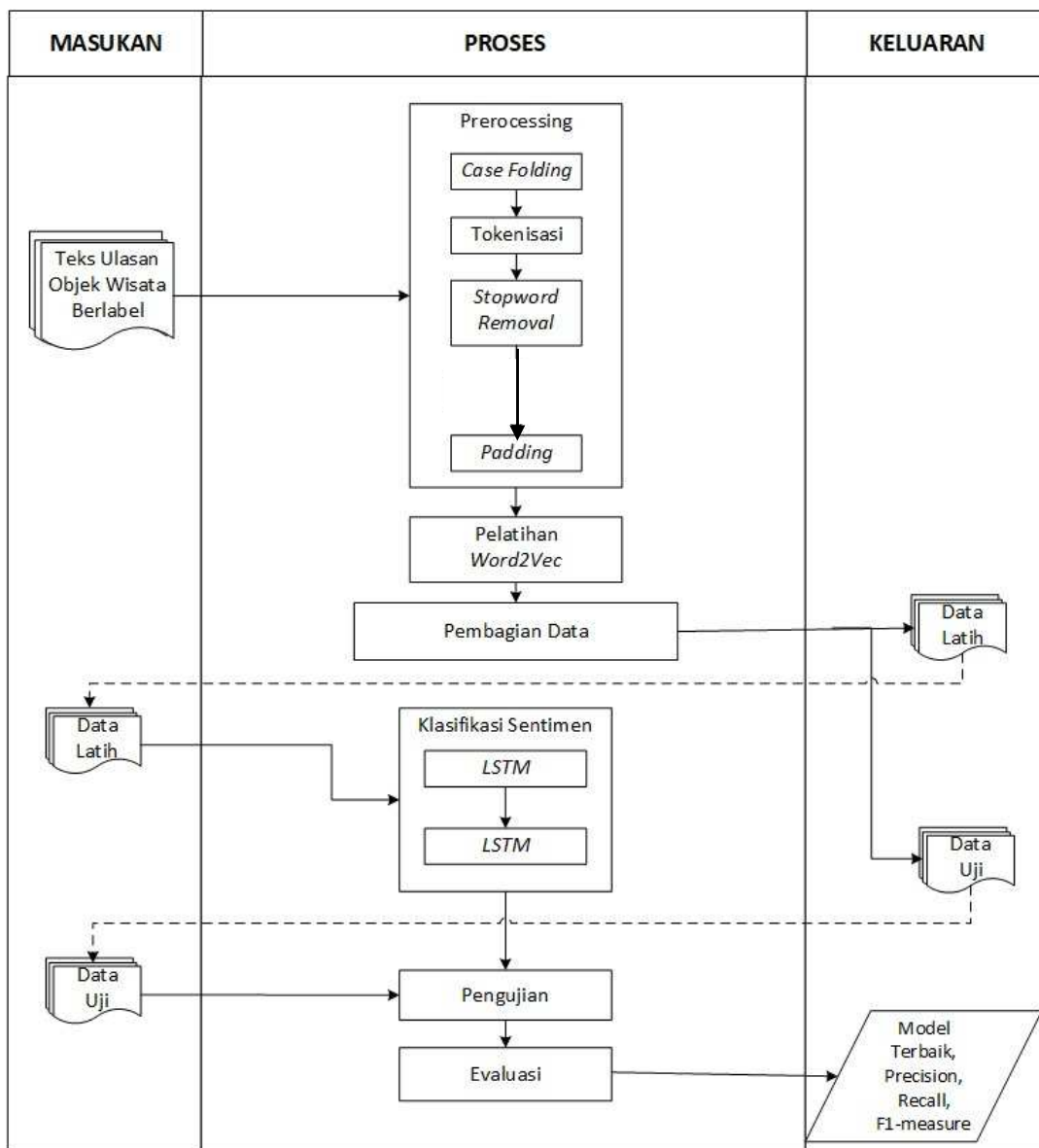
3.3.1 Pengumpulan Data

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 sentimen.

3.3.2 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 redundansi. 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 melalui Gambar 3.2.



Gambar 3.2 Alur Program

3.3.2 *Preprocessing*

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

1. *Case folding*

Case folding merupakan proses untuk mengubah semua karakter huruf pada sebuah kalimat menjadi huruf kecil atau huruf besar disebut *case folding*. *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*

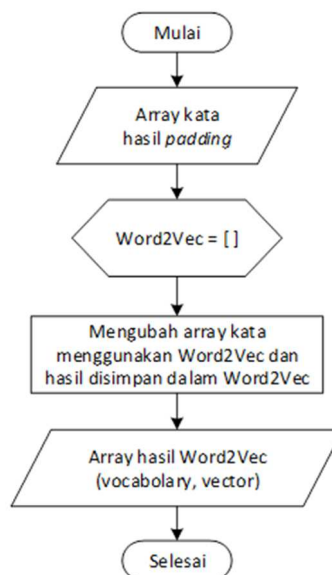
Stopword Removal merupakan tahap pengambilan kata-kata penting dan membuang kata-kata yang dianggap tidak penting. Cara untuk membuang kata yang tidak penting disebut *stopword removal*. *Stopword removal* bertujuan untuk menghilangkan kata-kata yang sering muncul namun tidak memiliki kontribusi dalam 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 *input* mempunyai panjang yang sama dengan cara menambahkan kata “<pad>”.

3.3.3 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 3.3.



Gambar 3.3 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:

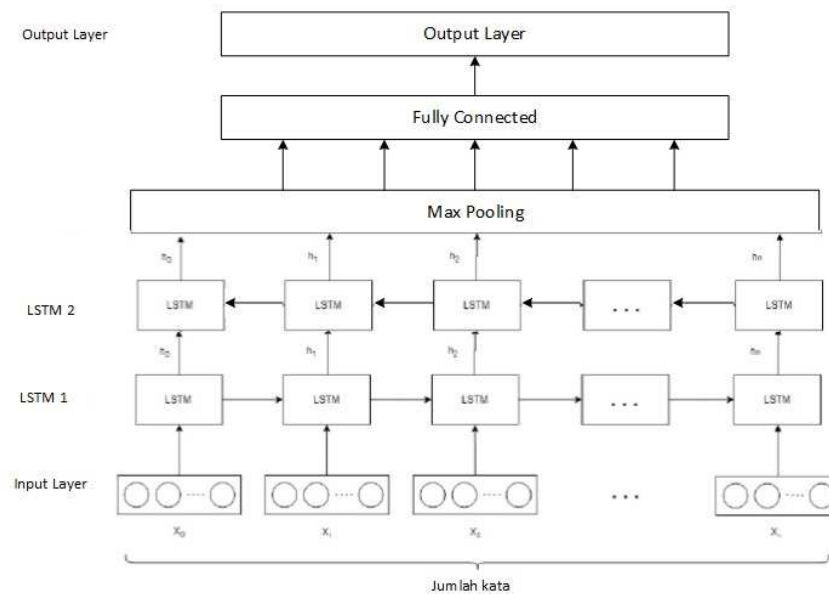
1. Arsitektur *Word2Vec*: *CBOW (Continuous Bag of Words)*, dan *Skip-gram-gram*
2. Metode evaluasi: *Hierarchical Softmax*, dan *Negative Sampling*
3. Ukuran dimensi: 100, 200, dan 300.

3.3.4 Pembagian Data

Pembagian data dilakukan dengan membagi data sebagai 80% sebagai data latih dan 20% sebagai data uji.

3.3.5 Pembentukan Model

1. Rancangan Arsitektur Model



Gambar 3.6 Rancangan Arsitektur Bi-LSTM dan *Word2Vec*

Rancangan arsitektur metode Bi-LSTM dapat dilihat pada Gambar 3.10. Berdasarkan Gambar 3.6 *input* dari arsitektur model Bi-LSTM adalah *array* hasil *preprocessing* teks dengan ukuran *JumlahKata*. 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.

3.3.6 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 3.2 menggambarkan contoh untuk prediksi dua kelas dengan menggunakan *confusion matrix* yang terdiri dari akurasi, *precision*, dan *recall*.

Tabel 3.2 Tabel *Confusion Matrix* pada Dua Kelas

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

Keterangan Tabel:

- *True Negative* jika hasil prediksi negatif dan data aktualnya negatif.
- *True Positive* jika hasil prediksi positif dan data aktualnya positif.
- *False Negative* jika hasil prediksi negatif dan data aktualnya positif.
- *False Positive* jika hasil prediksi positif dan data aktualnya negatif.

Terdapat beberapa persamaan yang telah ditetapkan pada matriks dua kelas yang memiliki persamaan seperti pada Persamaan 2.1 sampai Persamaan 2.3.

$$Accuracy = \frac{TP+FN}{TP+TN+FN+FP} \quad (3.1)$$

Keterangan Persamaan 2.1 sampai Persamaan 2.3:

- *Accuracy* merupakan proporsi kasus yang diidentifikasi benar dengan jumlah semua kasus.
- *Recall* merupakan proporsi kasus positif yang diidentifikasi dengan benar.
- *Precision* merupakan proporsi kasus dengan hasil positif yang benar.

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1 Hasil Penelitian

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.

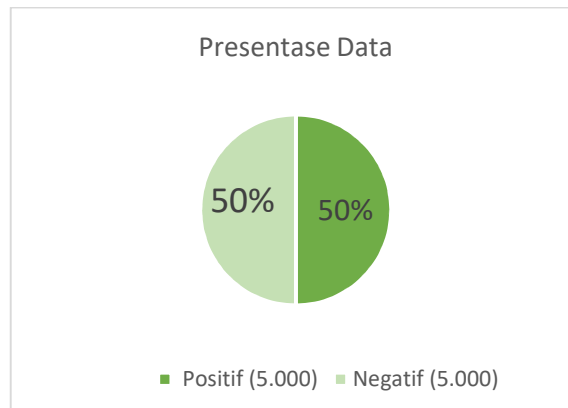
4.2 Pembahasan

4.4.1 Skenario Penelitian

Skenario penelitian membahas mengenai data pengujian dan skenario pengujian analisis sentimen menggunakan model *Word2Vec* dan model *Bidirectional Long Short Term Memory CNN*) untuk ulasan objek wisata di Pulau Bali yang berbahasa Indonesia.

1. Data Pengujian

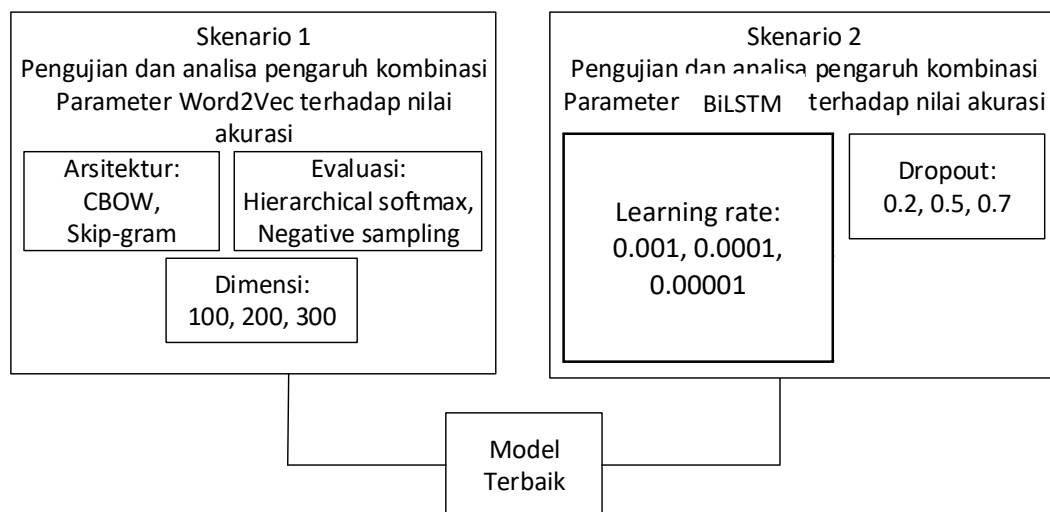
Ilustrasi pembagian data dapat dilihat pada Gambar 4.10. Penelitian ini menggunakan data sebanyak 10.000 data dimana data tersebut diambil dari total hasil *scrapping* data sebesar 15.100 data. Data yang digunakan pada penelitian ini 10.000 data berupa 5.000 data ulasan berlabel positif dan 5.000 data 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 4.1.



Gambar 0.1 Pembagian data pemodelan

2. Skenario Pengujian

Terdapat dua skenario pengujian pada penelitian ini. Skenario 1 merupakan pengujian dan analisa pengaruh kombinasi parameter *Word2Vec* terhadap nilai akurasi, sedangkan skenario 2 merupakan pengujian dan analisa pengaruh kombinasi parameter BiLSTM terhadap nilai akurasi. Gambaran umum skenario pengujian dijelaskan pada Gambar 4.11.



Gambar 0.21 Skenario pengujian

Terdapat tujuh parameter dalam penelitian ini yang digunakan untuk pengujian. Ketujuh parameter tersebut dibandingkan kinerjanya terhadap model *Bidirectional Long Short Term Memory* (Bi-LSTM). Kelima parameter tersebut terbagi menjadi 3 parameter untuk pengujian *Word2Vec* dan 2 parameter untuk

pengujian BiLSTM. 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 BiLSTM.

Pada model BiLSTM 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 LSTM-CNN 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:

a. Skenario 1

Skenario 1 berisi pengujian dan analisa terhadap pengaruh kombinasi dari model *Word2Vec*. Berikut penjelasan masing-masing parameter dari skenario 1:

- 1) Pengujian dilakukan untuk mengetahui pengaruh arsitektur *Word2Vec* dalam mendapatkan nilai akurasi terbaik. Arsitektur *Word2Vec* yang digunakan pada proses pengujian yaitu model CBOW dan model *Skip-gram*. Eksperimen yang terbentuk pada penelitian ini sebanyak 108, terdiri dari model CBOW sebanyak 54 dan model *Skip-gram* sebanyak 54. Data masukan pada skenario ini berupa data hasil prapengolahan teks.
- 2) 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.

- 3) Pengujian dilakukan untuk mengetahui pengaruh dimensi *Word2Vec* yang digunakan dalam mendapatkan nilai akurasi terbaik. Dimensi *Word2Vec* yang digunakan 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.

b. Skenario 2

Skenario 2 berisi pengujian dan analisa terhadap pengaruh kombinasi dari model LSTM-CNN. Berikut penjelasan masing-masing parameter dari skenario 2:

- 1) 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 ini berupa *array* hasil *Word2Vec* yang ada pada skenario 1.
- 2) 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.

2.4.2 Hasil dan Analisa Penelitian

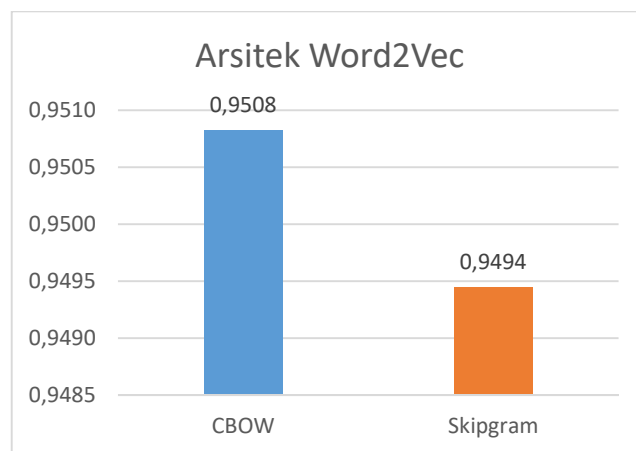
Sub-bab ini menjelaskan hasil dan analisa nilai akurasi dari skenario pengujian sebanyak 108 kombinasi. Model terbaik dipilih dengan menganalisa nilai akurasi untuk kombinasi *Word2Vec* berupa arsitektur, metode evaluasi, dan dimensi serta BiLSTM berupa *dropout*, dan *learning rate*. Model terbaik yang terpilih selanjutnya digunakan untuk analisis sentimen objek wisata.

1. Hasil dan analisa skenario 1

Pengujian skenario 1 dilakukan menggunakan kombinasi parameter *Word2Vec* dan parameter BiLSTM yang hasilnya dapat dilihat pada Lampiran 1. Hasil pada Lampiran 1 merupakan nilai akurasi menggunakan *confusion matrix* dari kombinasi 108 eksperimen. Pengaruh arsitektur, metode evaluasi, dan dimensi *Word2Vec* terhadap nilai akurasi dapat dilihat nilai akurasi yang dihitung untuk masing-masing eksperimen. Selanjutnya dihitung nilai rata-ratanya untuk setiap parameter arsitektur CBOW dan arsitektur *Skip-gram*, metode evaluasi *Hierarchical Softmax* dan *Negative Sampling*, dan dimensi 100, 200, 300.

a. Hasil dan Analisa Pengaruh Arsitektur *Word2Vec*

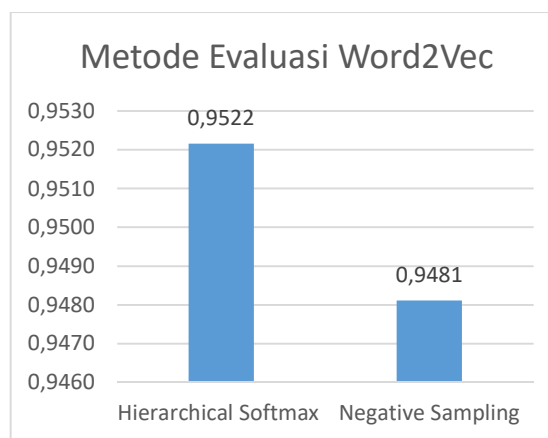
Grafik pada Gambar 4.3 menunjukkan arsitektur *Word2Vec Skip-gram* memiliki nilai rata-rata akurasi lebih baik dibandingkan dengan arsitektur *Word2Vec* CBOW. Akurasi rata-rata untuk *Skip-gram* sebesar 95,08%, sedangkan arsitektur dengan menggunakan CBOW memiliki rata-rata akurasi 94,94%. Model CBOW memproses kata konteks untuk memprediksi kata target, sedangkan model *Skip-gram* bekerja dengan cara memprediksi konteks yang diberikan dari satu kata. Umumnya pada analisis sentimen akan lebih optimal jika kata-kata yang target lebih menjadi perhatian. Oleh karena itu, arsitektur *Word2Vec* yang paling baik untuk model BiLSTM pada penelitian ini adalah CBOW.



Gambar 0.32 Perbandingan akurasi berdasarkan arsitektur *Word2Vec*

b. Hasil dan Analisa Pengaruh Metode Evaluasi *Word2Vec*

Grafik pada Gambar 4.4 menunjukkan metode evaluasi *Word2Vec* menggunakan *Hierarchical Softmax* memiliki akurasi yang lebih baik dibandingkan dengan metode evaluasi *Word2Vec* menggunakan *Negative Sampling*. Model *Word2Vec* dengan menggunakan metode evaluasi *Hierarchical Softmax* memiliki rata-rata akurasi 95,22%, sedangkan metode evaluasi *Negative Sampling* memiliki rata-rata akurasi 94,81%. Metode evaluasi *Negative Sampling* lebih sederhana dibandingkan *Hierarchical Softmax* karena hanya memperbarui sampel beberapa kata *output* sebagai sampel negatif. Metode evaluasi *Hierarchical Softmax* selama proses pelatihan menggunakan model *binary tree* untuk merepresentasikan semua katanya pada *vocabulary* dan simpul daun mewakili kata-kata yang jarang sehingga kata yang jarang muncul pasti akan mewarisi representasi vektor di atasnya. Oleh karena itu, metode evaluasi *Word2Vec* yang paling baik untuk model LSTM-CNN pada penelitian ini adalah *Hierarchical Softmax*.

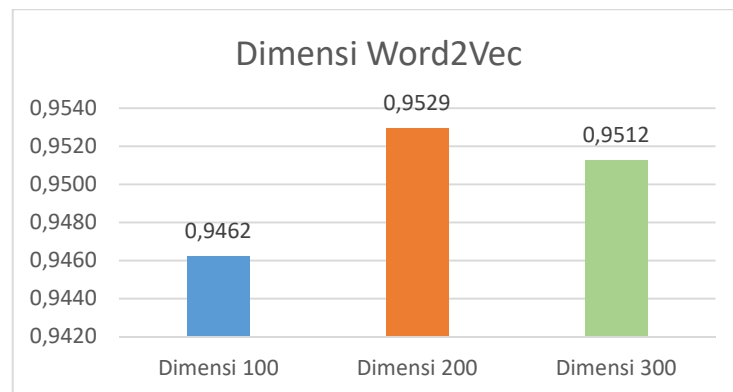


Gambar 0.4 Perbandingan akurasi berdasarkan metode evaluasi *Word2Vec*

c. Hasil dan Analisa Pengaruh Dimensi *Word2Vec*

Grafik pada Gambar 4.5 menunjukkan dimensi *Word2Vec* dengan dimensi bernilai 300 memiliki akurasi yang terbaik dibandingkan dengan dimensi *Word2Vec* bernilai 100 dan 200 Model *Word2Vec*. Dimensi 100 memiliki rata-rata akurasi 94,62%, sedangkan model *Word2Vec* dimensi 200 memiliki rata-rata akurasi 95,29% dan dimensi *Word2Vec* dengan nilai 300 memiliki rata-rata akurasi

95,12%. Dapat ditarik kesimpulan bahwa dimensi yang digunakan pada model *Word2Vec* mempengaruhi akurasi. Jika dimensi yang digunakan terlalu besar maka akurasinya kurang optimal, begitu juga apabila dimensinya terlalu kecil, akurasinya juga kurang optimal. Oleh karena itu, dimensi *Word2Vec* yang paling baik untuk model BiLSTM pada penelitian ini adalah 200.



Gambar 0.5 Perbandingan akurasi berdasarkan dimensi *Word2Vec*

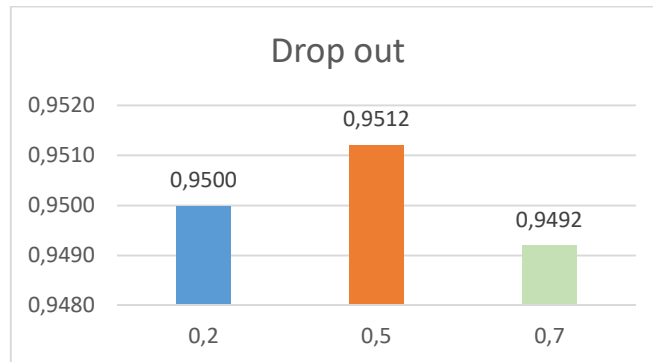
2. Hasil dan Analisa Skenario 2

Pengujian skenario 2 dilakukan menggunakan kombinasi parameter *Word2Vec* dan parameter LSTM yang hasilnya dapat dilihat pada Lampiran 1. Hasil pada Lampiran 1 merupakan nilai akurasi menggunakan *confusion matrix* dari kombinasi 108 eksperimen. Pengaruh *dropout*, dan *learning rate* terhadap nilai akurasi dapat dilihat dari nilai akurasi yang dihitung untuk masing-masing eksperimen. Selanjutnya dihitung nilai rata-ratanya untuk setiap parameter *dropout* bernilai 0,2; 0,5; 0,7, dan *learning rate* bernilai 0,001;0,0001; dan 0,00001.

a. Hasil dan Analisa Pengaruh *Dropout*

Grafik pada Gambar 4.6 menunjukkan *Dropout* menggunakan 0,7 memiliki akurasi yang lebih baik dibandingkan dengan *dropout* menggunakan 0,5 dan 0,2. *Dropout* menggunakan 0,7 memiliki rata-rata akurasi 95,00 %, sedangkan *dropout* menggunakan 0,5 dan 0,2 memiliki rata-rata akurasi 95,12% dan 94,92%. Dari hasil tersebut dapat disimpulkan *dropout* yang terlalu besar dan terlalu kecil akan

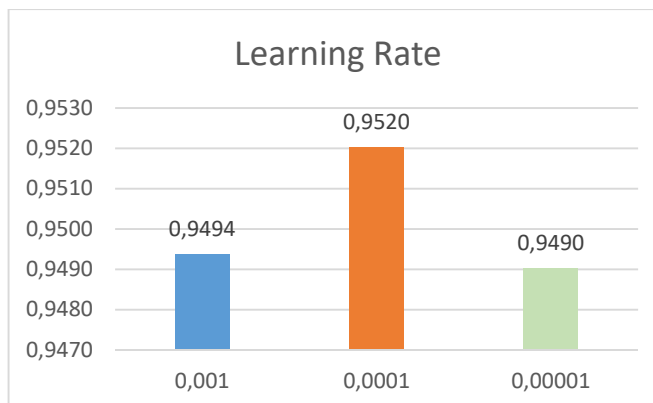
menjadikan akurasi tidak optimal. Oleh karena itu, *dropout* yang paling baik untuk model LSTM-CNN pada penelitian ini adalah 0,7.



Gambar 0.6 Perbandingan nilai akurasi berdasarkan *dropout*

b. Hasil dan Analisa Pengaruh *Learning Rate*

Grafik pada Gambar 4.7 menunjukkan *learning rate* menggunakan nilai 0,0001 memiliki akurasi yang lebih baik dibandingkan dengan *learning rate* dengan nilai 0,001 dan 0,00001. *Learning rate* menggunakan nilai 0,0001 memiliki rata-rata akurasi 95,20%, sedangkan *learning rate* menggunakan nilai 0,001 memiliki rata-rata akurasi 94,94%, serta *learning rate* menggunakan nilai 0,00001 memiliki rata-rata akurasi 94,90%. Dari hasil tersebut dapat disimpulkan bahwa *learning rate* yang terlalu besar dan terlalu kecil akan menjadikan akurasi tidak optimal. Oleh karena itu, *learning rate* yang paling baik untuk model LSTM-CNN pada penelitian ini adalah 0,0001.



Gambar 0.7 Perbandingan nilai akurasi berdasarkan *learning rate*

3. Hasil dan Analisa Seluruh Skenario

Hasil pengujian dari scenario 1 dan 2 adalah sebagai berikut:

1. CBOW dan Skipgram merupakan arsitektur *Word2Vec* yang diujikan. Hasil rata-rata dihitung untuk setiap arsitektur *Word2Vec* berdasarkan hasil yang ada pada Lampiran 1, arsitektur *Word2Vec* CBOW menghasilkan rata-rata akurasi yang lebih baik dibandingkan dengan arsitektur *Word2Vec* *Skipgram*.
2. *Hierarchical Softmax* dan *Negative Sampling* merupakan metode evaluasi *Word2Vec* yang diujikan. Hasil rata-rata dihitung untuk setiap metode evaluasi *Word2Vec* berdasarkan hasil yang ada pada Lampiran 1, metode evaluasi *Word2Vec* *Hierarchical Softmax* menghasilkan rata-rata akurasi yang 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* berdasarkan hasil yang ada pada Lampiran 1, 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* berdasarkan hasil yang ada pada Lampiran 1, *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* berdasarkan hasil yang ada pada Lampiran 1, *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 CBOV, metode evaluasi *Hierarchical Softmax*, dan dimensi 200, , *dropout* dengan nilai 0,5 , dan *learning rate* dengan nilai 0,0001. Kombinasi tersebut menghasilkan nilai akurasi tertinggi dari keseluruhan 108 kombinasi yang hasilnya dapat dilihat pada Lampiran 1, dengan nilai akurasi sebesar 96,86%.

4.3 Luaran yang Dicapai

Luaran yang di capai pada penelitian ini antara lain:

Luaran	Nama	Status
Model Terbaik	Model Terbaik BiLSTM terbaik untuk analisis sentiment ulasan wisata Pulau Bali	Sudah didapatkan
Jurnal	Jurnal Sinta 2	Draft
HKI	Model Terbaik BiLSTM terbaik untuk analisis sentiment ulasan wisata Pulau Bali	Draft
Bahan Buku Ajar	Bahan Buku Ajar Matakuliah Data Mining/ Big Data / Kecerdasan Buatan	Draft

BAB IV

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan dari hasil penelitian yang dilakukan mengenai analisis sentimen menggunakan model *Bidirectional Long Short Term Memory-Convolutional Neural Network* (BiLSTM) dan model *Word2Vec* terhadap ulasan objek wisata Pulau Bali di tripadvisor berbahasa Indonesia:

1. Model *Word2Vec* yang menghasilkan akurasi paling baik untuk melakukan klasifikasi dengan menggunakan parameter arsitektur CBOW, metode evaluasi *hierarchical softmax*, dan dimensi 200. Hal tersebut dikarenakan model CBOW memberikan akurasi yang baik untuk kata yang ditarget menjadi perhatian. Sedangkan metode evaluasi *Hierarchical Softmax* memberikan hasil yang lebih baik karena selama proses pelatihan menggunakan model *binary tree* untuk merepresentasikan semua katanya pada *vocabulary* dan simpul daun mewakili kata-kata yang jarang sehingga kata yang jarang muncul pasti akan mewarisi representasi vektor di atasnya. Selain itu, dimensi yang digunakan juga mempengaruhi nilai rata-rata akurasi karena dimensi yang terlalu kecil dan dimensi yang terlalu besar dapat mengurangi nilai akurasi.
2. Model BiLSTM menghasilkan akurasi terbaik untuk melakukan klasifikasi dengan menggunakan parameter *dropout 0,5*, dan *learning rate 0,0001*. *Dropout 0,5* lebih baik karena *neuron* yang terlalu banyak pada model LSTM-CNN justru akan mengurangi akurasi. *Learning rate 0,0001* lebih baik dibandingkan 0,001 dan 0,00001 karena *learning rate* yang terlalu besar atau terlalu kecil dapat menjadikan akurasi tidak optimal.
3. Kombinasi parameter model BiLSTM dan model *Word2Vec* yang terbaik menghasilkan nilai akurasi sebesar 96,86% dan digunakan untuk membuat klasifikasi analisis sentimen objek wisata berbahasa Indonesia.

5.2 Saran

Berdasarkan hasil akurasi BiLSTM dan *Word2Vec* terhadap analisis sentimen ulasan objek wisata Pulau Bali pada situs Tripadvisor.com berbahasa Indonesia yang cukup tinggi, maka saran untuk penelitian selanjutnya adalah melakukan *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.

DAFTAR PUSTAKA

- [1] M. Antara and M. S. Sumarniasih, "Role of Tourism in Economy of Bali and Indonesia," *J. Tour. Hosp. Manag.*, vol. 5, no. 2, pp. 34–44, 2017, doi: 10.15640/jthm.v5n2a4.
- [2] G. Sabou, "Influence of social media in choice of touristic destination Influence of social media in choice of touristic destination," vol. 3, no. February 2012, pp. 24–30, 2014.
- [3] B. Liu, "Sentiment analysis: Mining opinions, sentiments, and emotions," *Sentim. Anal. Min. Opin. Sentim. Emot.*, no. May, pp. 1–367, 2015, doi: 10.1017/CBO9781139084789.
- [4] A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of sentiment reviews using n-gram machine learning approach," *Expert Syst. Appl.*, vol. 57, pp. 117–126, 2016, doi: 10.1016/j.eswa.2016.03.028.
- [5] A. Yenter and A. Verma, "Deep CNN-LSTM with combined kernels from multiple branches for IMDb review sentiment analysis," 2017 IEEE 8th Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2017, vol. 2018-Janua, pp. 540–546, 2017, doi: 10.1109/UEMCON.2017.8249013.
- [6] H. Ghulam, F. Zeng, W. Li, and Y. Xiao, "Deep Learning-Based Sentiment Analysis for Roman Urdu Text," *Procedia Comput. Sci.*, vol. 147, pp. 131–135, 2019, doi: 10.1016/j.procs.2019.01.202.
- [7] A. Aziz Sharfuddin, M. Nafis Tihami, and M. Saiful Islam, "A Deep Recurrent Neural Network with BiLSTM model for Sentiment Classification," 2018 Int. Conf. Bangla Speech Lang. Process. ICBSLP 2018, pp. 1–4, 2018, doi: 10.1109/ICBSLP.2018.8554396.
- [8] K. Zhang, W. Song, L. Liu, X. Zhao, and C. Du, "Bidirectional long short-term memory for sentiment analysis of Chinese product reviews," *ICEIEC 2019 - Proc. 2019 IEEE 9th Int. Conf. Electron. Inf. Emerg. Commun.*, pp. 665–668, 2019, doi: 10.1109/ICEIEC.2019.8784560.
- [9] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment analysis of comment

- texts based on BiLSTM,” *IEEE Access*, vol. 7, no. c, pp. 51522–51532, 2019, doi: 10.1109/ACCESS.2019.2909919.
- [10]R. P. Nawangsari, R. Kusumaningrum, and A. Wibowo, “Word2vec for Indonesian sentiment analysis towards hotel reviews: An evaluation study,” *Procedia Comput. Sci.*, vol. 157, pp. 360–366, 2019, doi: 10.1016/j.procs.2019.08.178.
- [11]D. I. Af’idah, R. Kusumaningrum, and B. Surarso, “Long short term memory convolutional neural network for Indonesian sentiment analysis towards touristic destination reviews,” *Proc. - 2020 Int. Semin. Appl. Technol. Inf. Commun. IT Challenges Sustain. Scalability, Secur. Age Digit. Disruption, iSemantic 2020*, pp. 630–637, 2020, doi: 10.1109/iSemantic50169.2020.9234210.
- [12]D. Li and J. Qian, “Text sentiment analysis based on long short-term memory,” *2016 1st IEEE Int. Conf. Comput. Commun. Internet, ICCCI 2016*, pp. 471–475, 2016, doi: 10.1109/CCI.2016.7778967.
- [13]Jiddy Abdillah, Ibnu Asror, and Yanuar Firdaus Arie Wibowo, “Emotion Classification of Song Lyrics using Bidirectional LSTM Method with GloVe Word Representation Weighting,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 4, pp. 723–729, 2020, doi: 10.29207/resti.v4i4.2156.

Lampiran 1.

NO	W2V ID	Model	Metode Evaluasi	Dimensi	DO	LR	Akurasi
BiLSTM-001	W2V -01	CBOW	HS	100	0,2	0,001	0.9287
BiLSTM-002	W2V -01	CBOW	HS	100	0,2	0,0001	0.9577
BiLSTM-003	W2V -01	CBOW	HS	100	0,2	0,00001	0.9531
BiLSTM-004	W2V -01	CBOW	HS	100	0,5	0,001	0.9331
BiLSTM-005	W2V -01	CBOW	HS	100	0,5	0,0001	0.9588
BiLSTM-006	W2V -01	CBOW	HS	100	0,5	0,00001	0.9630
BiLSTM-007	W2V -01	CBOW	HS	100	0,7	0,001	0.9550
BiLSTM-008	W2V -01	CBOW	HS	100	0,7	0,0001	0.9625
BiLSTM-009	W2V -01	CBOW	HS	100	0,7	0,00001	0.9467
BiLSTM-010	W2V -02	CBOW	HS	200	0,2	0,001	0.9601
BiLSTM-011	W2V -02	CBOW	HS	200	0,2	0,0001	0.9593
BiLSTM-012	W2V -02	CBOW	HS	200	0,2	0,00001	0.9561
BiLSTM-013	W2V -02	CBOW	HS	200	0,5	0,001	0.9543
BiLSTM-014	W2V -02	CBOW	HS	200	0,5	0,0001	0.9565
BiLSTM-015	W2V -02	CBOW	HS	200	0,5	0,00001	0.9686
BiLSTM-016	W2V -02	CBOW	HS	200	0,7	0,001	0.9613
BiLSTM-017	W2V -02	CBOW	HS	200	0,7	0,0001	0.9552
BiLSTM-018	W2V -02	CBOW	HS	200	0,7	0,00001	0.9591
BiLSTM-019	W2V -03	CBOW	HS	300	0,2	0,001	0.9412
BiLSTM-020	W2V -03	CBOW	HS	300	0,2	0,0001	0.9569
BiLSTM-021	W2V -03	CBOW	HS	300	0,2	0,00001	0.9570
BiLSTM-022	W2V -03	CBOW	HS	300	0,5	0,001	0.9522
BiLSTM-023	W2V -03	CBOW	HS	300	0,5	0,0001	0.9539
BiLSTM-024	W2V -03	CBOW	HS	300	0,5	0,00001	0.9503
BiLSTM-025	W2V -03	CBOW	HS	300	0,7	0,001	0.9467
BiLSTM-026	W2V -03	CBOW	HS	300	0,7	0,0001	0.9606
BiLSTM-027	W2V -03	CBOW	HS	300	0,7	0,00001	0.9494
BiLSTM-028	W2V -04	CBOW	NS	100	0,2	0,001	0.9597
BiLSTM-029	W2V -04	CBOW	NS	100	0,2	0,0001	0.9539
BiLSTM-030	W2V -04	CBOW	NS	100	0,2	0,00001	0.9550
BiLSTM-031	W2V -04	CBOW	NS	100	0,5	0,001	0.9526
BiLSTM-032	W2V -04	CBOW	NS	100	0,5	0,0001	0.9427
BiLSTM-033	W2V -04	CBOW	NS	100	0,5	0,00001	0.9449
BiLSTM-034	W2V -04	CBOW	NS	100	0,7	0,001	0.9502
BiLSTM-035	W2V -04	CBOW	NS	100	0,7	0,0001	0.9516
BiLSTM-036	W2V -04	CBOW	NS	100	0,7	0,00001	0.8731
BiLSTM-037	W2V -05	CBOW	NS	200	0,2	0,001	0.9488
BiLSTM-038	W2V -05	CBOW	NS	200	0,2	0,0001	0.9508
BiLSTM-039	W2V -05	CBOW	NS	200	0,2	0,00001	0.9490
BiLSTM-040	W2V -05	CBOW	NS	200	0,5	0,001	0.9488
BiLSTM-041	W2V -05	CBOW	NS	200	0,5	0,0001	0.9532
BiLSTM-042	W2V -05	CBOW	NS	200	0,5	0,00001	0.9239
BiLSTM-043	W2V -05	CBOW	NS	200	0,7	0,001	0.9604

BiLSTM-044	W2V -05	CBOW	NS	200	0,7	0,0001	0.9543
BiLSTM-045	W2V -05	CBOW	NS	200	0,7	0,00001	0.9561
BiLSTM-046	W2V -06	CBOW	NS	300	0,2	0,001	0.9441
BiLSTM-047	W2V -06	CBOW	NS	300	0,2	0,0001	0.9612
BiLSTM-048	W2V -06	CBOW	NS	300	0,2	0,00001	0.9450
BiLSTM-049	W2V -06	CBOW	NS	300	0,5	0,001	0.9226
BiLSTM-050	W2V -06	CBOW	NS	300	0,5	0,0001	0.9615
BiLSTM-051	W2V -06	CBOW	NS	300	0,5	0,00001	0.9535
BiLSTM-052	W2V -06	CBOW	NS	300	0,7	0,001	0.9606
BiLSTM-053	W2V -06	CBOW	NS	300	0,7	0,0001	0.9559
BiLSTM-054	W2V -06	CBOW	NS	300	0,7	0,00001	0.9538
BiLSTM-055	W2V -07	SKIPGRAM	HS	100	0,2	0,001	0.9483
BiLSTM-056	W2V -07	SKIPGRAM	HS	100	0,2	0,0001	0.9485
BiLSTM-057	W2V -07	SKIPGRAM	HS	100	0,2	0,00001	0.9146
BiLSTM-058	W2V -07	SKIPGRAM	HS	100	0,5	0,001	0.9449
BiLSTM-059	W2V -07	SKIPGRAM	HS	100	0,5	0,0001	0.9201
BiLSTM-060	W2V -07	SKIPGRAM	HS	100	0,5	0,00001	0.9515
BiLSTM-061	W2V -07	SKIPGRAM	HS	100	0,7	0,001	0.9439
BiLSTM-062	W2V -07	SKIPGRAM	HS	100	0,7	0,0001	0.9660
BiLSTM-063	W2V -07	SKIPGRAM	HS	100	0,7	0,00001	0.9560
BiLSTM-064	W2V -08	SKIPGRAM	HS	200	0,2	0,001	0.9496
BiLSTM-065	W2V -08	SKIPGRAM	HS	200	0,2	0,0001	0.9474
BiLSTM-066	W2V -08	SKIPGRAM	HS	200	0,2	0,00001	0.9492
BiLSTM-067	W2V -08	SKIPGRAM	HS	200	0,5	0,001	0.9524
BiLSTM-068	W2V -08	SKIPGRAM	HS	200	0,5	0,0001	0.9560
BiLSTM-069	W2V -08	SKIPGRAM	HS	200	0,5	0,00001	0.9587
BiLSTM-070	W2V -08	SKIPGRAM	HS	200	0,7	0,001	0.9422
BiLSTM-071	W2V -08	SKIPGRAM	HS	200	0,7	0,0001	0.9486
BiLSTM-072	W2V -08	SKIPGRAM	HS	200	0,7	0,00001	0.9598
BiLSTM-073	W2V -09	SKIPGRAM	HS	300	0,2	0,001	0.9569
BiLSTM-074	W2V -09	SKIPGRAM	HS	300	0,2	0,0001	0.9483
BiLSTM-075	W2V -09	SKIPGRAM	HS	300	0,2	0,00001	0.9597
BiLSTM-076	W2V -09	SKIPGRAM	HS	300	0,5	0,001	0.9504
BiLSTM-077	W2V -09	SKIPGRAM	HS	300	0,5	0,0001	0.9655
BiLSTM-078	W2V -09	SKIPGRAM	HS	300	0,5	0,00001	0.9550
BiLSTM-079	W2V -09	SKIPGRAM	HS	300	0,7	0,001	0.9528
BiLSTM-080	W2V -09	SKIPGRAM	HS	300	0,7	0,0001	0.9533
BiLSTM-081	W2V -09	SKIPGRAM	HS	300	0,7	0,00001	0.9598
BiLSTM-082	W2V -10	SKIPGRAM	NS	100	0,2	0,001	0.9450
BiLSTM-083	W2V -10	SKIPGRAM	NS	100	0,2	0,0001	0.9419
BiLSTM-084	W2V -10	SKIPGRAM	NS	100	0,2	0,00001	0.9478
BiLSTM-085	W2V -10	SKIPGRAM	NS	100	0,5	0,001	0.9504
BiLSTM-086	W2V -10	SKIPGRAM	NS	100	0,5	0,0001	0.9469
BiLSTM-087	W2V -10	SKIPGRAM	NS	100	0,5	0,00001	0.9535
BiLSTM-088	W2V -10	SKIPGRAM	NS	100	0,7	0,001	0.9595
BiLSTM-089	W2V -10	SKIPGRAM	NS	100	0,7	0,0001	0.9353

BiLSTM-090	W2V -10	SKIPGRAM	NS	100	0,7	0,00001	0.9483
BiLSTM-091	W2V -11	SKIPGRAM	NS	200	0,2	0,001	0.9338
BiLSTM-092	W2V -11	SKIPGRAM	NS	200	0,2	0,0001	0.9518
BiLSTM-093	W2V -11	SKIPGRAM	NS	200	0,2	0,00001	0.9506
BiLSTM-094	W2V -11	SKIPGRAM	NS	200	0,5	0,001	0.9536
BiLSTM-095	W2V -11	SKIPGRAM	NS	200	0,5	0,0001	0.9687
BiLSTM-096	W2V -11	SKIPGRAM	NS	200	0,5	0,00001	0.9463
BiLSTM-097	W2V -11	SKIPGRAM	NS	200	0,7	0,001	0.9563
BiLSTM-098	W2V -11	SKIPGRAM	NS	200	0,7	0,0001	0.9584
BiLSTM-099	W2V -11	SKIPGRAM	NS	200	0,7	0,00001	0.9460
BiLSTM-100	W2V -12	SKIPGRAM	NS	300	0,2	0,001	0.9573
BiLSTM-101	W2V -12	SKIPGRAM	NS	300	0,2	0,0001	0.9599
BiLSTM-102	W2V -12	SKIPGRAM	NS	300	0,2	0,00001	0.9516
BiLSTM-103	W2V -12	SKIPGRAM	NS	300	0,5	0,001	0.9499
BiLSTM-104	W2V -12	SKIPGRAM	NS	300	0,5	0,0001	0.9034
BiLSTM-105	W2V -12	SKIPGRAM	NS	300	0,5	0,00001	0.9496
BiLSTM-106	W2V -12	SKIPGRAM	NS	300	0,7	0,001	0.9496
BiLSTM-107	W2V -12	SKIPGRAM	NS	300	0,7	0,0001	0.9460
BiLSTM-108	W2V -12	SKIPGRAM	NS	300	0,7	0,00001	0.9494

Lampiran 2.

ORGANISASI PENGUSUL PENELITIAN

1. Ketua

Nama	: Dwi Intan Af'idah, M.Kom
NIPY	: 11.020.470
NIDN	: 0620089203
Pangkat/Golongan	: III/b
Jabatan Fungsional	: -
Jabatan Struktural	: Koordinator Akademik
Bidang Ilmu	: Sistem Informasi
Pengalaman Penelitian	: -

2. Anggota 1

Nama	: Dairoh, M.Sc
NIPY	: 04.14.178
NIDN	: 0612108701
Pangkat/Golongan	: IIIC
Jabatan Fungsional	: Lektor
Jabatan Struktural	: -
Bidang Ilmu	: Fisika Komputasi
Pengalaman Penelitian	:

- 1) Sistem informasi manajemen pusat penelitian dan pengabdian kepada masyarakat (studi kasus : pusat penelitian dan pengabdian kepada masyarakat (p3m))
- 2) Implementasi kecerdasan buatan dalam pembuatan robot kontrol dengan fuzzy logic
- 3) *Text mining* berbasis feature weights untuk penilaian sentiment wisatawan terhadap tempat tujuan wisata kota tegal sebagai upaya pengembangan potensi daerah bidang pariwisata
- 4) Sistem Informasi Geografis Pertanian di Kabupaten Brebes
- 5) Implementasi wavelet pada data seismik gunungapi
- 6) Dekomposisi Wavelet Data Seismik Broadband dari Stasiun Wanagama Yogyakarta pada saat Letusan Gunung Merapi 2010
- 7) Opinion Mining Terhadap Toko Online Di Media Sosial Menggunakan Algoritma Naïve Bayes (Studi Kasus: Akun Facebook Dugal Delivery)
- 8) Sistem Informasi Pelacakan Jejak Alumni (Tracer Study) Berbasis Smart Phone Pada Politeknik Harapan Bersama
- 9) Segmentasi Dan Perbaikan Citra Untuk Proses Pengukuran Dimensi Beras
- 10) Implementasi Wavelet Pada Data Seismik Gunungapi
- 11) Text Mining berbasis Feature Weights untuk Penilaian sentiment Wisatawan terhadap tempat tujuan wisata kota Tegal sebagai Upaya Pengembangan Potensi Daerah bidang pariwisata

- 12) Evaluasi Ukuran Partikel dan Efek Farmakologi Sediaan Footh sanitizer Spray Kombinasi Ekstrak Biji Kopi Dan Rimpang jahe
- 13) Model Optimal Neural Network Sebagai Penunjang Keputusan Antipasi Pra Lulusan Agar Bekerja Sesuai Lulusan
- 14) Rancang Bangun Pengembangan Sistem Informasi Manajemen Penelitian Dan Pengabdian Masyarakat Politeknik Harapan Bersama (Simpn Di PHB)

3. Anggota 2

Nama	: Susi Nurindah Sari
NIM	: 19090102
Semester	: Empat
Pengalaman Penelitian	: -

Lampiran 3.**JUSTIFIKASI PENGGUNAAN ANGGARAN PENELITIAN**

1. Pengembangan Model					
No.	Justifikasi Pemakaian				Jumlah Honor
1	Pengumpulan Data				Rp 500.000
2	Analisis Data (olah data)				Rp 500.000
3	Pembuatan Model				Rp 500.000
Sub Total (Rp)					Rp. 1.500.000
2. Bahan Habis Pakai dan Peralatan					
No	Material	Justifikasi Pemakaian	Satuan	Harga Satuan	Jumlah Harga
1	Material 1	Tinta Print	1 Buah	Rp 200.000	Rp 200.000
2	Material 2	Fotocopy dan Penjilidan	5 Jilid	Rp 20.000	Rp 100.000
3	Material 3	Kerta A4 berat 80 gram	1 Rim	Rp 57.000	Rp 57.000
4	Konsumsi 1	Snack + makan siang	12 kali	Rp 25.000	Rp 300.000
5	Publikasi	Save tee Jurnal sinta 2 jurnal Matriks	1	Rp 1.000.000	Rp 1.000.00
Sub Total (Rp)					Rp. 1.657.000
3. Penunjang					
No.	Justifikasi Pemakaian		Satuan	Harga Satuan	Jumlah Harga
1	pembuatan +Jilid Proposal dan Laporan		6	Rp 50.000	Rp 300.000
Sub Total (Rp)					Rp 300.000
Jumlah Total					Rp 3.457.000

Tegal, Agustus 2021

Mengetahui,

Ketua P3M

Ketua Peneliti

Kusnadi, M.Pd.
NIPY. 014.015.217

Intan Dwi Af'idah, S.T., M.Kom.
NIPY. 11.020.470

Lampiran 4.

SUSUNAN ORGANISASI TIM PELAKSANA PENELITIAN DAN PEMBAGIAN TUGAS

No	Nama/NIDN	Instansi Asal	Bidang Ilmu	Alokasi Waktu (Jam/Minggu)	Uraian Tugas
1	Dwi Intan Af'idah, M.Kom/ 0620089203	Politeknik Harapan Bersama	Sistem Informasi	48 Jam	<ol style="list-style-type: none">a. Mengkoordinasi proses pengambilan data, pengumpulan data, analisis data, penyusunan interpretasi data.b. Mengkoordinasi persiapan instrumen penelitian, perlengkapan penelitian, dan instrumen penunjang.c. Bertanggungjawab dalam mendesain model, dan pengembangan model.d. Mengkoordinasi penyusunan laporan akhir penelitian, publikasi hasil penelitian dalam bentuk model terbaik, prosiding internasional scopus IOP Universitas Muhammadiyah Magelang dan buku Ajar.e. Bertanggung jawab terhadap hasil pelaporan penelitian mulai seminar hasil laporan akhir dan penggunaan anggaran penelitian
2	Dairoh, M.Sc/ 0612108701	Politeknik Harapan Bersama	Fisika Komputasi	48 Jam	<ol style="list-style-type: none">a. Turut bertanggung jawab dalam proses pengambilan data, pengumpulan data, analisis data,

					penyusunan interpretasi data. b. Turut bertanggung jawab terhadap hasil pelaporan penelitian mulai dari seminar laporan akhir dan penggunaan anggaran penelitian
--	--	--	--	--	---

Lampiran 5



Yayasan Pendidikan Harapan Bersama
PoliTeknik Harapan Bersama
Kampus I : Jl. Mataram No.9 Tegal 52142 Telp. 0283-352000 Fax. 0283-353353
Kampus II : Jl. Dewi Sartika No. 71 Tegal 52117 Telp. 0283-350567
Website : www.politektegal.ac.id | Email : sekretariat@politektegal.ac.id

**SURAT KEPUTUSAN
DIREKTUR POLITEKNIK HARAPAN BERSAMA
NOMOR: 098.05/PHB/V/2021**

**TENTANG
PENERIMA PENDANAAN HIBAH KOMPETITIF PENELITIAN DAN
PENGABDIAN MASYARAKAT OLEH INSTITUSI
BAGI DOSEN POLITEKNIK HARAPAN BERSAMA
TAHUN ANGGARAN 2020/2021 SEMESTER GENAP**

- DIREKTUR POLITEKNIK HARAPAN BERSAMA,**
- Menimbang** :
- bahwa untuk meningkatkan kualitas dan kuantitas pelaksanaan penelitian dan pengabdian kepada masyarakat bagi Dosen di Politeknik Harapan Bersama, maka perlu menetapkan kebijakan dalam bidang pendanaan penelitian dan pengabdian kepada masyarakat;
 - bahwa untuk tertib administrasi keuangan dalam pendanaan penelitian dan pengabdian kepada masyarakat, maka perlu ditetapkan tahapan penyerahan pendanaan oleh institusi untuk hibah kompetitif penelitian dan pengabdian masyarakat kepada Dosen Politeknik harapan Bersama;
 - bahwa nama-nama yang tercantum dalam lampiran telah lolos kualifikasi untuk menerima pendanaan hibah kompetitif dari Institusi;
 - berdasarkan pertimbangan sebagaimana dimaksud pada huruf a dan b, dipandang perlu menetapkan Surat Keputusan Direktur Politeknik Harapan Bersama;
- Mengingat** :
- Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional (Lembaran Negara Republik Indonesia Tahun 2003 Nomor 78, Tambahan Lembaran Negara Republik Indonesia Tahun 2003 Nomor 4301);
 - Undang-Undang Republik Indonesia Nomor 28 Tahun 2004 tentang Perubahan Undang-Undang Nomor 16 Tahun 2001 tentang Yayasan (Lembaran Negara Republik Indonesia Tahun 2004 Nomor 115, Tambahan Lembaran Negara Republik Indonesia Tahun 2004 Nomor 4430);
 - Undang-Undang Republik Indonesia Nomor 14 Tahun 2005 tentang Guru dan Dosen (Lembaran Negara Republik Indonesia Tahun 2005 Nomor 157, Tambahan Lembaran Negara Republik Indonesia Tahun 2005 Nomor 4586);
 - Undang-Undang Republik Indonesia Nomor 12 Tahun 2012 tentang Pendidikan Tinggi (Lembaran Negara Republik Indonesia Tahun 2012 Nomor 158, Tambahan Lembaran Negara Republik Indonesia Tahun 2012 Nomor 5336);
 - Peraturan Pemerintah..

-
5. Peraturan Pemerintah Republik Indonesia Nomor 4 Tahun 2014 tentang Penyelenggaraan Pendidikan Tinggi dan Pengelolaan Perguruan Tinggi (Lembaran Negara Republik Indonesia Tahun 2014 Nomor 16, Tambahan Lembaran Negara Republik Indonesia Tahun 2014 Nomor 5500);
6. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 3 Tahun 2020 tentang Standar Nasional Pendidikan Tinggi (Berita Negara Republik Indonesia Tahun 2020 Nomor 47);
7. Keputusan Menteri Pendidikan Nasional Republik Indonesia Nomor: 128/D/0/2002 tentang Pemberian Ijin Penyelenggaraan Program-Program Studi dan Pendirian Politeknik Harapan Bersama di Tegal yang diselenggarakan oleh Yayasan Pendidikan Harapan Bersama di Tegal;
8. Keputusan Menteri Hukum dan Hak Asasi Manusia Republik Indonesia Nomor: AHU-2674.AH.01.04 Tahun 2012 tentang pengesahan Yayasan Pendidikan Harapan Bersama (Tambahan Berita Negara Republik Indonesia Tanggal 20/6-2014 No. 49);
9. Keputusan Menteri Riset, Teknologi, dan Pendidikan Tinggi Nomor: 231/KPT/I/2018 tentang Yayasan Pendidikan Harapan Bersama sebagai Badan Penyelenggara Politeknik Harapan Bersama;
10. Surat Keputusan Yayasan Pendidikan Harapan Bersama Nomor 114.05/YPHB/XII/2020 tentang Statuta Politeknik Harapan Bersama;
- Memperhatikan : Surat Pemberitahuan Pusat Penelitian dan Pengabdian Masyarakat (P3M) Nomor: 064.03/P3M.PHB/III/2021 tentang pengajuan dan penerimaan proposal Penelitian dan Pengabdian Kepada Masyarakat Politeknik Harapan Bersama Semester Genap Tahun Akademik 2020/2021.

MEMUTUSKAN:

- Menetapkan : Surat Keputusan Direktur Politeknik Harapan Bersama tentang Penerima Pendanaan Oleh Institusi Untuk Hibah Kompetitif Penelitian dan Pengabdian Masyarakat Bagi Dosen Politeknik Harapan Bersama Tahun Anggaran 2020/2021.
- Pertama : Menetapkan nama yang tercantum dalam lampiran Keputusan ini sebagai Penerima Pendanaan Oleh Institusi Untuk Hibah Kompetitif Penelitian dan Pengabdian Masyarakat Bagi Dosen Politeknik Harapan Bersama Tahun Anggaran 2020/2021.
- Kedua : 1. Pemberian bantuan dana penelitian minimal Rp. 2.000.000,- (Dua juta rupiah) per judul;
2. Pemberian bantuan dana pengabdian kepada masyarakat minimal Rp. 2.000.000,- (Dua juta rupiah) per judul);
3. Pembayaran dilakukan dengan 2 (dua) tahap, yaitu:
a. Pembayaran tahap I sebesar 60% dari total dana yang didapatkan setelah menyerahkan proposal dan perjanjian yang telah ditandatangani oleh Direktur Politeknik Harapan Bersama;
b. Pembayaran Tahap II sebesar 30% dari total dana yang didapatkan setelah menyerahkan laporan hasil; dan
c. 10% dari total dana yang didapatkan diserahkan kepada P3M.

- Ketiga : Dosen yang melaksanakan Penelitian dan/atau Pengabdian Kepada Masyarakat wajib menyerahkan laporan hasil kepada Direktur dan Wakil Direktur I melalui Pusat Penelitian dan Pengabdian Masyarakat (P3M), meliputi:
- Laporan penelitian sebanyak 2 (dua) eksemplar;
 - Softcopy Jurnal;
 - Softcopy.
- Keempat : Semua produk hasil penelitian dan pengabdian masyarakat termasuk Paten menjadi hak milik Politeknik Harapan Bersama.
- Kelima : Surat Keputusan ini berlaku sejak tanggal ditetapkan dan apabila di kemudian hari terdapat kekeliruan akan diadakan perbaikan sebagaimana mestinya.

Ditetapkan di: Tegal
Pada tanggal: 31 Mei 2021
Direktur,

Nizar Suhendra, S.E., M.P.P.
NIPY.08.020.008

Lampiran: Surat Keputusan Direktur Politeknik Harapan Bersama
 Tentang : Penerima Pendanaan Oleh Institusi Untuk Hibah Kompetitif Penelitian dan Pengabdian Masyarakat Bagi Dosen Politeknik Harapan Bersama Tahun Anggaran 2020/2021 Semester Genap
 Nomor : 096 .05/PHB/V/2021
 Tanggal : 31 Mei 2021

Daftar Penerima Bantuan Biaya Penelitian dan Pengabdian Kepada Masyarakat Bagi Dosen Program Studi Sarjana Terapan Politeknik Harapan Bersama Tahun Akademik 2020/2021 Semester Genap

NO	KETUA	JUDUL	PRODI	SKEMA	NOMINAL
1	Ghea Dwi Rahmadiane, S.E., M.Si. Arifany Ferida, S.E., M.Si. Nizar Fahrezi Majid	Upaya Inovasi Dan Kreativitas Pelaku UMKM Kota Tegal Di Masa Pandemi Covid-19	Sarjana Terapan Akuntansi Sektor Publik	Penelitian	Rp. 3,271,000
2	Nurul Mahmudah, S.E., M. Si., AK, CA Yusri Anis Faidah, S.E., M. Si. Laelatul Maghfiroh	Transparansi Dan Akuntabilitas Dalam Pengelolaan Anggaran Pendapatan Dan Belanja Desa (Apdes) (Studi Kasus: Desa Banjaranyar Kecamatan Brebes Kabupaten Brebes)	Sarjana Terapan Akuntansi Sektor Publik	Penelitian	Rp. 3,228,500
3	Dwi Intan Afidah, S.T., M. Kom, Dairoh, M.Sc. Susi Nurindah sari	<i>Bidirectional Long Short Term Memory</i> Dan <i>Word2vec</i> Untuk Analisis Sentimen Terhadap Ulasan Destinasi Wisata Pulau Bali	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 3,457,000
4	Dega Surono Wibowo, ST, M.Kom. Ardi Susanto, S.Kom., M. Cs. Khibar Pusaka	Analisis Dan Pengujian Celah Keamanan Pada Website DIV Teknik Informatika Politeknik Harapan Bersama	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 2,814,000
5	Sharfina Febbi Handayani M.Kom, Riszi Wijayatun Pratiwi, S.Kom., M.Cs. Mulyana Putriyani	Analisis Sentimen Pada Data Ulasan <i>Twitter</i> Dengan Menggunakan <i>Long Short Term Memory</i>	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 3,342,500
6	Dyah Apriliani, ST, M.Kom. Hepatika Zidny Ilmadina, S. Pd., M. Kom, Nurlaela	Sentiment Analysis Penilaian Toko Online Menggunakan <i>Naive Bayes</i> Dan <i>Neural Network</i>	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 3,314,000
7	Taufiq Abidin, S.Pd, M. Kom. Slamet Wiyono, S.Pd., M. Eng. Agung Iswanto.	Implementasi Algoritma Nrf Dalam <i>Recommender System</i> Berbasis <i>Content</i> Dan <i>Collaborative Filtering</i> Sebagai Strategi Bisnis UMKM	Sarjana Terapan Teknik Informatika	Penelitian	Rp. 3,242,500

Lampiran: Surat Keputusan Direktur Politeknik
Harapan Bersama

Tentang : Penerima Pendanaan Oleh Institusi
Untuk Hibah Kompetitif Penelitian dan
Pengabdian Masyarakat Bagi Dosen
Politeknik Harapan Bersama Tahun
Anggaran 2020/2021 Semester Genap

Nomor : 098.05/PHB/V/2021

Tanggal : 31 Mei 2021

13	Slamet Wiyono, S. Pd., M. Eng Dega Surono Wibowo, S.T., M. Kom, Riszi Wijayatsun Pratiwi, S.Kom., M.Cs. Naimatul Maulidiyah Getar Dewantara Agung Iswanto	Pemanfaatan Teknik <i>Scraping</i> Data Untuk Perencanaan Usaha Jualan Online Menggunakan <i>Marketplace</i>	Sarjana Terapan Teknik Informatika	PKM	Rp. 2,900,000
14	Muhammad Fikri Hidayattullah, S.T., M.Kom. Dega Surono Wibowo, S.T., M. Kom, Ardi Susanto, S.Kom., M.Cs. Alfin Auzikri Wildan Sania Alfiansyah	Pengenalan <i>Software</i> Al- Mausu'ah Al-Hadits Bagi Santri Madrasah Fiqih Sumber Ilmu Dalam Melakukan Studi Takhrij Hadits	Sarjana Terapan Teknik Informatika	PKM	Rp. 2,787,500

Direktur,


Nizar Suhendra, S.E., MPP
NIPY.08.020.008

Bidirectional Long Short Term Memory Dan Word2vec Untuk Analisis Sentimen Terhadap Ulasan Destinasi Wisata Pulau Bali

Dwi Intan Afidah, Dairoh, Sharfina Febbi Handayani, Rizki Wijayatun Pratiwi, Susi Nurindah Sari
 Politeknik Harapan Bersama

Article Info

Article history:

Received, xxx xx xxxx
 Revised, xxx xx xxxx
 Accepted, xxx xx xxx

Kata Kunci:

Analisis Sentimen,
 Bidirectional Long Short Term
 Memory
 Word2Vec

ABSTRAK

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. Hal tersebut mengakibatkan perlu dilakukannya identifikasi objek wisata Pulau Bali yang diminati masyarakat dengan cara menganalisis ulasan masyarakat yang tersebar di internet mengenai suatu objek wisata. Penelitian ini bertujuan melakukan analisis sentimen terhadap ulasan objek wisata di Pulau Bali menggunakan metode Bi-LSTM (*Bidirectional Long Short Term Memory*) dan *Word2Vec*, sehingga diperoleh model terbaik dengan mengetahui kinerja dari kombinasi parameter Bi-LSTM dan *Word2Vec*. Bi-LSTM merupakan metode *deep learning* yang menawarkan akurasi yang lebih baik daripada metode LSTM biasa. Sedangkan *Word2Vec* sebagai salah satu metode *pretraining* dipilih karena dapat menangkap makna semantik teks dengan baik. Data yang akan digunakan dari penelitian ini berupa data ulasan objek wisata pulau Bali yang berasal dari situs tripadvisor.com. Rencana kegiatan pada penelitian ini antara lain, pengumpulan data, perancangan alur program, *preprocessing*, *pretraining Word2Vec*, pembagian data menjadi data uji dan latih, pembentukan model dengan proses pelatihan dan pengujian, serta evaluasi hasil pengujian untuk menentukan model terbaik.

ABSTRACT

Keyword:

Proper management of tourism on the island of Bali, especially in new tourism objects that have the potential to increase the number of tourist arrivals, to

Skenario 2 berisi pengujian dan analisa terhadap pengaruh kombinasi dari model LSTM-CNN. Berikut penjelasan masing-masing parameter dari skenario 2:

- 1) 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 ini berupa *array* hasil *Word2Vec* yang ada pada skenario 1.
- 2) Pengaruh nilai *learning rate* yang digunakan untuk mendapatkan nilai akurasi terbaik. Nilai *learning rate* yang digunakan pada proses pengujian sebesar 0,001; 0,0001; 0,00001. 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.

4. KESIMPULAN

Kesimpulan dari hasil penelitian yang dilakukan mengenai analisis sentimen menggunakan model Bidirectional Long Short Term Memory-Convolutional Neural Network (BiLSTM) dan model Word2Vec terhadap ulasan objek wisata Pulau Bali di tripadvisor berbahasa Indonesia:

- a. Model Word2Vec yang menghasilkan akurasi paling baik untuk melakukan klasifikasi dengan menggunakan parameter *uvectors*, CBOW, metode evaluasi *hierarchical softmax*, dan dimensi 200. Hal tersebut dikarenakan model CBOW memberikan akurasi yang baik untuk kata yang ditarget menjadi perhatian. Sedangkan metode evaluasi *Hierarchical Softmax* memberikan hasil yang lebih baik karena selama proses pelatihan menggunakan model *binary tree* untuk merepresentasikan semua kata pada *vocabulary* dan simpul daun mewakili kata-kata yang jarang sehingga kata yang jarang muncul pasti akan mewarisi representasi vektor di atasnya. Selain itu, dimensi yang digunakan juga mempengaruhi nilai rata-rata akurasi karena dimensi yang terlalu kecil dan dimensi yang terlalu besar dapat mengurangi nilai akurasi.
- b. Model BiLSTM menghasilkan akurasi terbaik untuk melakukan klasifikasi dengan menggunakan parameter *dropout* 0,5, dan *learning rate* 0,0001. *Dropout* 0,5 lebih baik karena neuron yang terlalu banyak pada model LSTM-CNN justru akan mengurangi akurasi. *Learning rate* 0,0001 lebih baik dibandingkan 0,001 dan 0,00001 karena *learning rate* yang terlalu besar atau terlalu kecil dapat menjadikan akurasi tidak optimal.
- c. Kombinasi parameter model BiLSTM dan model Word2Vec yang terbaik menghasilkan nilai akurasi sebesar 96,86% dan digunakan untuk membuat klasifikasi analisis sentimen objek wisata berbahasa Indonesia.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terimakasih untuk P3M Politeknik Harapan Bersama yang memberikan fasilitas dalam pelaksanaan penelitian ini

REFERENSI

- [1] M. Antara and M. S. Sumamiasih, "Role of Tourism in Economy of Bali and Indonesia," *J. Tour. Hosp. Manag.*, vol. 5, no. 2, pp. 34–44, 2017, doi: 10.15640/jhm.v5i2a4.
- [2] G. Sabou, "Influence of social media in choice of touristic destination Influence of social media in choice of touristic destination," vol. 3, no. February 2012, pp. 24–30, 2014.
- [3] B. Liu, "Sentiment analysis: Mining opinions, sentiments, and emotions," *Sentim. Anal. Min. Opin. Sentim. Emot.*, no. May, pp. 1–367, 2015, doi: 10.1017/CBO9781139084789.
- [4] A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of sentiment reviews using n-gram machine learning

UEMCON 2017, vol. 2018-Janua, pp. 540–546, 2017, doi: 10.1109/UEMCON.2017.8249013.

- [6] H. Ghulam, F. Zeng, W. Li, and Y. Xiao, "Deep Learning-Based Sentiment Analysis for Roman Urdu Text," *Procedia Comput. Sci.*, vol. 147, pp. 131–135, 2019, doi: 10.1016/j.procs.2019.01.202.
- [7] A. Aziz Shariffuddin, M. Nafis Tihani, and M. Saiful Islam, "A Deep Recurrent Neural Network with BiLSTM model for Sentiment Classification," 2018 Int. Conf. Bangla Speech Lang. Process. ICBSLP 2018, pp. 1–4, 2018, doi: 10.1109/ICBSLP.2018.8554396.
- [8] K. Zhang, W. Song, L. Liu, X. Zhao, and C. Da, "Bidirectional long short-term memory for sentiment analysis of Chinese product reviews," *ICEIEC 2019 - Proc. 2019 IEEE 9th Int. Conf. Electron. Inf. Emerg. Commun.*, pp. 665–668, 2019, doi: 10.1109/ICEIEC.2019.8784560.
- [9] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment analysis of comment texts based on BiLSTM," *IEEE Access*, vol. 7, no. c, pp. 51522–51532, 2019, doi: 10.1109/ACCESS.2019.2909919.
- [10] R. P. Nawangsari, R. Kusumaningrum, and A. Wibowo, "Word2vec for Indonesian sentiment analysis towards hotel reviews: An evaluation study," *Procedia Comput. Sci.*, vol. 157, pp. 360–366, 2019, doi: 10.1016/j.procs.2019.08.178.
- [11] D. I. Al'Idah, R. Kusumaningrum, and B. Sutarsa, "Long short term memory convolutional neural network for Indonesian sentiment analysis towards touristic destination reviews," *Proc. - 2020 Int. Semin. Appl. Technol. Inf. Commun. IT Challenges Sustain. Scalability, Secur. Age Digit. Disruption, iSemantic 2020*, pp. 630–637, 2020, doi: 10.1109/iSemantic50169.2020.9234210.
- [12] D. Li and J. Qian, "Text sentiment analysis based on long short-term memory," 2016 1st IEEE Int. Conf. Comput. Commun. Internet, ICCCI 2016, pp. 471–475, 2016, doi: 10.1109/CCL2016.7778967.
- [13] Jidky Abdillah, Ibnu Astor, and Yamar Firdaus Arie Wibowo, "Emotion Classification of Song Lyrics using Bidirectional LSTM Method with GloVe Word Representation Weighting," *J. RESTI (Rekayasa Sist. dan Teknol. Informatika)*, vol. 4, no. 4, pp. 723–729, 2020, doi: 10.29207/resti.v4i4.2156.