**Ekstraksi Aspek pada ulasan restoran menggunakan *domain specific word embedding***

**Tugas Akhir**

**diajukan untuk memenuhi salah satu syarat**

**memperoleh gelar sarjana**

**dari Program Studi S1 Informatika**

**Fakultas Informatika**

**Universitas Telkom**

**1301170012**

**Ahmad Satriamulya**

****

**Program Studi Sarjana Informatika**

**Fakultas Informatika**

**Universitas Telkom**

**Bandung**

**2022**

LEMBAR PENGESAHAN

**Ekstraksi Aspek pada ulasan restoran menggunakan *domain specific word embedding***

***Aspect Extraction on restaurant reviews using domain-specific word embedding***

**NIM : 1301170012**

**Ahmad Satriamulya**

Tugas akhir ini telah diterima dan disahkan untuk memenuhi sebagian syarat memperoleh gelar pada Program Studi Sarjana Informatika  
Fakultas Informatika

Universitas Telkom

Bandung, 05/04/2022

Menyetujui

|  |  |  |
| --- | --- | --- |
| Pembimbing I, |  | Ketua Program Studi  Sarjana Informatika, |
|  |  |  |
| Dr. Ade Romadhony, S.T., M.T  NIP: 06840042 |  | Dr. Erwin Budi Setiawan, S.Si., M.T.  NIP: 00760045 |

|  |
| --- |
|  |

LEMBAR PERNYATAAN

Dengan ini saya, Ahmad Satriamulya, menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul Ekstraksi Aspek pada ulasan restoran menggunakan domain specific word embedding beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang belaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika di kemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam buku TA atau jika ada klaim dari pihak lain terhadap keaslian karya,

Bandung, 5 April 2022

Yang Menyatakan

Ahmad Satriamulya

Ekstraksi Aspek pada ulasan restoran menggunakan domain specific word embedding

Ahmad Satriamulya1, Dr. Ade Romadhony, S.T., M.T 2

1,2,3Fakultas Informatika, Universitas Telkom, Bandung

4Divisi Digital Service PT Telekomunikasi Indonesia

1 [ahmadrayi@student.telkomuniversity.ac.id,](mailto:ahmadrayi@student.telkomuniversity.ac.id,), 2 [aderomadhony@telkomuniversity.ac.id](mailto:aderomadhony@telkomuniversity.ac.id).

#### Abstrak

**Ulasan di internet bisa menjadi bagian penting untuk sebuah bisnis dan dapat memengaruhi pemilik ataupun konsumen dalam membuat keputusan. Dengan mudahnya akses informasi dengan bantuan internet dan gawai, informasi berupa pendapat, pengalaman, dan umpan balik orang lain dapat digunakan sebagai acuan untuk melakukan sebuah tindakan. Untuk bisnis pada sektor makanan dan minuman konsumen biasa memberikan ulasan dengan sentimen negatif ataupun positif berdasarkan beberapa aspek dari bisnis terkait. Rasa makanan, suasana, harga, pelayanan merupakan contoh aspek yang umum dituliskan dalam sebuah ulasan. Pada Tugas Akhir ini, akan dilakukan ekstraksi aspek untuk teks ulasan konsumen terhadap restoran yang berada di indonesia. Metode yang digunakan adalah Deep Learning dengan arsitektur Recurrent Neural Network, menggunakan Domain Specific Embedding sebagai Word Embedding, dan beberapa upaya untuk mengurangi out of vocabulary pada model.**

**Kata kunci : ekstraksi aspek, Recurrent Neural Network, word embedding, out-of-vocabulary word**

# **Abstract**

**Reviews on the internet can be an important part of a business and can influence owners or consumers in making decisions. With easy access to information with the help of the internet and devices, information in the form of opinions, experiences, and feedback from others can be used as a reference for taking an action. For businesses in the food and beverage sector, consumers usually provide reviews with negative or positive sentiments based on several aspects of the related business. The taste of food, atmosphere, price, service are examples of aspects that are commonly written in a review. In this final project, aspect extraction will be carried out for the text of consumer reviews of restaurants in Indonesia. The method used is Deep Learning with Recurrent Neural Network architecture, using Domain Specific Embedding as Word Embedding, and several attempts to reduce out of vocabulary in the model.**

**Keywords: aspect extraction, Recurrent Neural Network, word embedding, out-of-vocabulary word**

1. **Pendahuluan**

**Latar Belakang**

Ulasan di internet bisa menjadi bagian penting untuk sebuah bisnis dan dapat memengaruhi pemilik ataupun konsumen dalam membuat keputusan. Dengan mudahnya akses informasi dengan bantuan internet dan gawai, informasi berupa pendapat, pengalaman, dan umpan balik orang lain dapat digunakan sebagai acuan untuk melakukan sebuah Tindakan [1]. Pada bisnis berbidang makan dan minuman, kepuasan konsumen dapat menggambarkan kualitas restoran, dan untuk konsumen, ulasan di internet dapat menjadi sumber informasi yang membantu untuk mengurangi risiko pembelian yang tidak diinginkan yang telah dirasakan oleh konsumen lain [2]. Pemilik bisnis juga bisa menggunakan opini konsumen untuk mengetahui lebih dalam mengenai perspektif konsumen terhadap bisnisnya sehingga bisa mengembangkan bisnisnya sesuai dengan keinginan konsumen-konsumennya [4]

Pada umumnya, konsumen memberikan ulasan terhadap restoran berupa opini yang dapat menggambarkan aspek-aspek dari sebuah restoran, contohnya makanan, harga, tempat, suasana, dan pelayanan. Ulasan tersebut bisa diklasifikasikan ke dalam dua kategori, yaitu positif dan negatif klasifikasi ini biasa disebut dengan sentimen [3].

Sentiment analysis merupakan bidang studi yang mancakup Analisa opini, sentimen, evaluasi, penilaian, sikap, dan emosi orang dari sebuah teks terhadap entitas contohnya produk, layanan, organisasi, individu. Sentiment analysis memiliki beberapa beberapa tingkatan antara lain, document level yang menganalisa sentimen secara keseluruhan teks, sentence level yang menganalisa sentimen pada setiap kalimat, dan aspect level yang menganalisa sentimen langsung berdasarkan opini yang terdapat dalam teks [3]. Sentiment analysis berbasis pada aspect level (aspect-based sentiment analysis) dapat memberikan analisis yang lebih kompleks dibandingkan kedua level lainnya [4]. Proses penting dalam aspect-based sentiment analysis adalah ekstraksi aspek dari teks dan membangun hubungan antara aspek tersebut dengan opini yang dimaksud [5].

Pada Tugas Akhir ini, akan dilakukan ekstraksi aspek dari ulasan restoran yang berada di Indonesia. Pada penilitian sebelumnya [6], ekstraksi fitur pada teks menggunakan tri-gram untuk mendapatkan konteks dari kata sebelum dan setelahnya, dan word2vec untuk membuat word embedding dari data Wikipedia dan menggunakan beberapa metode Deep Learning. Pada penelitian ini penulis akan melakukan penelitian dari saran penulis sebelumnya [6] yaitu meminimasi out-of-vocabulary (OOV) pada word embedding dengan cara menggunakan domain specific embedding dan beberapa teknik word pre-processing. Untuk domain specific word embedding akan dibuat dari data ulasan review restoran di Indonesia dan untuk word pre-processing akan dilakukan .

**Topik dan Batasannya**

Topik yang dibahas dalam tugas akhir ini adalah bagaimana cara menurunkan out-of-vocabulary pada permasalahan aspek ekstraksi pada domain review restoran menggunakan model LSTM dan Bi-LSTM serta mengukur performansi dari metode yang digunakan.

Batasan masalah pada tugas akhir ini adalah sebagai berikut: Pertama, data yang digunakan merupakan data review restoran yang telah dilabeli pada penlitian sebelumnya [6]. Kedua, label aspek terbagi kedalam 5 label yaitu food, price, ambience, service, dan miscellaneous.

**Tujuan dan Manfaat**

Tujuan dari tugas akhir ini adalah melakukan membandingkan metode-metode untuk mengurangi out-of-vocabulary word pada word embedding yang akan digunakan untuk ekstraksi aspek pada ulasan berdomain restoran serta mengukur performansi dari masing-masing metode. Mengenai manfaat dari penelitian ini, hasil ujian bisa digunakan sebagai referensi performa berbagai metode untuk mengurangi out-of-vocabulary word pada word embedding terutama pada permasalahan ekstraksi aspek dari ulasan restoran.

**Organisasi Tulisan**

Struktur penulisan dari tugas akhir ini disusun sebagai berikut: Bagian pertama berisi pendahuluan terkait tugas akhir ini. Bagian kedua menjelaskan studi yang terkait dengan tugas akhir ini. Bagian ketiga menjelaskan pemodelan dari sistem yang dibangun dan data yang digunakan. Bagian keempat menjelaskan hasil dan evaluasi hasil pengujian yang telah dilakukan pada bagian ketiga. Kemudian, pada bagian terakhir menjelaskan kesimpulan dan saran berdasarkan hasil pengujian yang dilakukan pada tugas akhir ini.

1. **Studi Terkait**

Pada riset sebelumnya yang dilakukan Widhianto, Rachmansyah Adhi, dan Ade Romadhony [6] melakukan ekstraksi aspek pada data ulasan restoran yang berada di Bandung, Indonesia dan membandingkan beberapa metode Machine Learning dan Deep Learning. Model RNN (LSTM dan BiLSTM) merupakan dua model terbaik dibandingkan dengan model SVM dan CNN dengan f1 score 79,1 pada 11 tag, dan 55,1 pada 10 tag ketika tidak menghitung tag O pada dataset dikarenakan mayoritas tagnya merupakan O. Pada riset tersebut, representasi kata menggunakan skip-gram dari Word2Vec yang dilatih dengan teks artikel wikipedia. Data train dilakukan dengan mengubah kalimat menjadi tri-gram dengan kata ke-2 sebagai acuan untuk output, contohnya trigram [rasanya, enak, sekali] akan memiliki aspek FOOD dikarenakan kata “enak” memiliki aspek FOOD.

R. Perera, T. Malepathirana, et al. [18] melakukan penelitian untuk meningkatkan hasil aspect aggregation pada review berdomain laptop dan restoran dan membandingkan hasil antara menggunakan domain independent word embedding dan domain specific embedding. Domain Independent embedding dibuat dengan dataset Google News corpus Word2Vec model dengan dimensi 300, domain specific word embedding yang dibuat dengan dataset review dari beberapa website dengan total data 31 juta kalimat untuk domain restoran dan 2 juta untuk laptop dan dibuat dengan skip-gram dengan vector seize 300, negative sample 5 dan window size 2. Pearsons Correlation yang didapat dengan menggunakan domain specific embedding adalah 0.616 pada domain laptop dan 0.541 pada domain restoran, sementara pada domain independent adalah 0.443 pada domain laptop dan 0.558 pada restoran.

Riset yang berkaitan dengan Bahasa sehari-hari atau tidak baku (*colloquial*) telah dilakukan oleh N. Aliyah Salsabila, et al. [10]. Pada riset tersebut dibuat data lexicon colloquial berisi 3,592 kata unique out of vocabulary beserta kata formalnya. Dataset tersebut dibuat dari komen figur publik Instagram [20] sebanyak 24,602 dan dianotasi secara manual. Data lexicon colloquial beserta kata formalnya dapat digunakan untuk menormalisasi sebuah teks dan setelah dilakukan percobaan untuk mendeteksi spam komen di Instagram dan ternyata tidak memiliki perbedaan yang signifkan. Hanafiah, Novita, et al. [19] melakukan nomalisasi teks pada twitter dengan kategori complain. Pada riset ini dibuat sebuah model normalisasi out of vocabulary untuk menormalkan kata yang tidak dapat dikenali. Model tersebut dibuat dengan kumpulan kata-kata basic, kamus kata informal, dan kamus konteks. Model ini mendapatkan akurasi 89.83% dalam mengubah kata out-of-vocabulary untuk menjadi kata yang dikenali.

**Ekstraksi Aspek**

Ekstraksi aspek adalah tugas untuk mengidentifikasi dan mengekstraksi aspek yang relevan yang biasa digunakan untuk penambangan opini (opinion mining) dan sentimen analisis, contoh aspeknya adalah atribut atau fitur dari produk [1]. Untuk setiap kata pada suatu teks, akan diberikan label aspek yang sesuai untuk teks tersebut, contohnya pada kalimat ulasan “kemarin makan disini enak dan suasananya nyaman”, kata “makan” sampai “enak” berkaitan dengan makanan sehingga dilabel aspek “FOOD” yang berarti berkaitan dengan makanan, sementara kata “suasana” dan “nyaman” dilabel aspek “AMBIENCE” karena berkaitan dengan suasana restoran. Pada tugas akhir ini, akan digunakan 5 aspek yaitu FOOD, PRICE, AMBIENCE, SERVICE, MISCELLANEUS, dan O ketika suatu kata tidak memiliki aspek yang telah dibuat. Label aspek tersebut digunakan sesuai dengan penelitian sebelumnya [6].

**Word Embedding**

Word embedding (representasi kata) adalah proses konversi kata yang berupa karakter alphanumeric kedalam bentuk vector. Tujuannya adalah membuat kata-kata yang memiliki properti tertentu, misalnya berada pada konteks yang sama, atau memiliki arti semantik yang sama tidak memiliki jarak yang jauh.

Word2Vec merupakan salah satu metode untuk membuat word embedding. Word2Vec membuat word embedding dengan arsitektur neural network yang memiliki 1 hidden [22]. Terdapat 2 metode yang sering digunakan dalam pembuatan word embedding melalui Word2Vec yaitu skip-gram dan CBOW. Pada metode skip-gram, neural network akan berusaha memprediksi kata2 konteks berdasarkan kata utamanya, sementara CBOW sebaliknya, memprediksi kata utama dari konteks kata2 yang diberikan. Contohnya pada kalimat “kemarin makan disini enak dan suasananya nyaman”, ketika kata utama yang akan dibuat word embeddingnya adalah “enak” maka jika parameter window = 2, konteks katanya adalah “makan”, “disini”, “dan”, “suasananya” tanpa menghiraukan jarak antara kata utama dengan konteks katanya. Model skip-gram akan bekerja lebih baik dengan data yang kecil dan dapat merepresentasikan kata dengan frekuensi yang sedikit, sementara CBOW dapat dilatih lebih cepat dan dapat merepresentasikan kata lebih baik jika frekuensi katanya banyak [23]. Pada tugas akhir ini, akan digunakan model Word2Vec dengan metode Skip-Gram, dikarenakan dataset yang relatif sedikit dan banyak kata-kata yang frekuensinya sedikit dikarenakan kata yang tidak baku pada ulasan berdomain restoran.

**Bi-LSTM**

Long short term memory (LSTM) adalah salah satu modifikasi dari recurrent neural network atau RNN. LSTM dibuat untuk melengkapi kekurangan RNN yang tidak dapat memprediksi kata berdasarkan informasi lampau yang disimpan dalam jangka waktu lama. Dengan demikian, LSTM dapat mengingat kumpulan informasi yang telah disimpan dalam jangka waktu panjang, dan juga dapat menghapus informasi yang tidak lagi relevan [24]. Dengan kemampuan LSTM untuk mengumpulkan dan membuang informasi yang lampau, LSTM sering digunakan pada permasalahan yang berbentuk sequensial dan membutuhkan konteks data sebelumnya seperti Natural Language Processing (NLP), Speech / handwriting recognition, Music generation.

Bi-Directional LSTM (BiLSTM) adalah modifikasi dari LSTM dengan memberikan arah balik sehingga bukan hanya memory dari belakang saja, tetapi dari depan juga. BiLSTM telah terbukti dapat memberikan peningkatan performa pada domain NLP [26, 27] dibandingkan dengan model LSTM, salah satu alasannya adalah konteks dari sebuah kalimat tidak hanya dilihat dari belakang saja, tapi bisa juga dari depan kalimat [25].

1. **Sistem yang Dibangun**

Sistem dibangun menggunakan bahasa pemrograman Python dengan alur seperti pada gambar 2.

******

**Dataset**

Dataset untuk training dan test model berupa teks ulasan restoran yang terdapat pada website Zomato dan berada di bandung dan terdapat total 4.798 ulasan restoran dengan rata-rata 75 kata per ulasan. Dataset ini telah dilabel pada penelitian sebelumnya [6] oleh Rachmansyah Adhi Widhianto dengan format label IOB2.

Untuk membuat domain word embedding, dataset ulasan yang digunakan merupakan ulasan restoran yang berada di Jakarta dan Bandung yang didapatkan melalui crawling website pergikuliner [9]. Pada dataset ini terdapat 88.115 total review dengan Panjang rata2 105 kata per ulasan.

**Pre-Processing**

Pre-processing merupakan focus utama penulis untuk meminimalisir jumlah out-of-vocabulary pada feature extraction (pembuatan word embedding). Semua metode pre-processing dilakukan dengan python dengan bermacam-macam library dan sumber, berikut merupakan metode-metode pre-processing yang akan digunakan untuk penilitan ini:

1. **Basic Text Preprocessing**, metode ini meliputi:
   1. Case folding, proses merubah huruf kapital menjadi huruf kecil.
   2. Punctuation removal, proses menghilangkan tanda baca
   3. Emotes & Emoji removal, proses menghilangkan emoji atau emotes
   4. Remove Extra whitespace, proses menghilangkan spasi berlebih
2. **Stemming**, metode yang digunakan untuk mengubah sebuah kata menjadi kata dasarnya. Pada penilitian ini akan digunakan python library stemmer berbahasa indonesia yaitu StemmerFactory dari Sastrawi.
3. **Alay Dictionary**, Metode ini bertujuan untuk menormalisasi kata yang sering ditulis pada ulasan restoran yang tidak baku menjadi lebih baku. Alay Pre-processing akan mengubah sebuah kata yang bisa tidak baku seperti “kane” yang memiliki makna “enak” dibalik, ataupun penggunaan huruf berlebih pada kata seperti “enaaaak” yang sebenarnya memiliki makna yang sama dengan kata “enak”. Untuk melakukan proses ini, digunakan beberapa sumber data dictionary:
   1. colloquial-indonesian-lexicon yang merupakan hasil riset oleh N. Aliyah Salsabila, Y [9]
   2. informal-formal-Indonesian-dictionary.tsv, yang merupakan hasil riset dari W Wibowo, Haryo Akbarianto, et al. [11]
4. **Rule-Based Processing**, Proses ini dilakukan karena ketika pengecekan out-of-vocabulary pada data secara manual, ditemukan bahwa terdapat beberapa kata out-of-vocabulary yang memiliki pattern yang sama, sehingga untuk menguranginya akan dibuat rule-based text processing yang dibagi menjadi 3 bagian, yaitu:
   1. **Normalisasi huruf berulang**, metode ini digunakan untuk mengurangi jumlah huruf berulang pada sebuah kata. Sebuah kata huruf yang sama berurutan lebih dari 2 kali seperti kata “kacaaaauuu” akan dikurangi semua huruf berulangnya menjadi maksimal 2 huruf berurut, sehingga kata tersebut akan menjadi “kacaauu”. Tahap ini sangat berpengaruh terutama ketika metode ke-3 (Alay Dictionary) pada penelitian ini juga dilakukan.
   2. **Normalisasi kata ulang**, metode ini digunakan untuk menyamakan format kata ulang. Contohnya pada kata “enakenak” (dimana tanda baca sudah terhapus pada Basic Text Processing) akan diubah menjadi enak2, contoh lainnya adalah “kemanamananya” menjadi “kemana2nya”. Proses ini dilakukan dengan Regular Expression (regex) dengan menyamakan kata yang muncul 2 kali secara berurutan dengan minimal 3 huruf dan merubah kemunculan kata ke-2 menjadi angka 2. Pemberian minimal 3 huruf ditetapkan agar beberapa kata seperti “bubur” tidak ikut berubah menjadi “bu2r”.
   3. **Normalisasi harga**, dikarenakan pada penlitian ini memiliki aspek “price” yang berhubungan dengan harga. Penulis mengimplementasikan normalisasi harga bertujuan untuk mempermudah model dalam mengenali harga. Beberapa regex digunakan untuk mengenali pattern harga yang sering muncul pada ulasan restoran. Berikut merupakan regex yang digunakan:
      1. Memiliki awalan “idr” atau “rp” (prefix) diikuti dengan digit
      2. berawalan digit diikuti dengan akhiran “k”, ”rb”, ”rupiah”, ”ribu”, ”rban”, atau ”an” (suffix)
      3. memiliki prefix diikuti dengan digit dan diakhiri dengan suffix
      4. 4 digit berurutan atau lebih

**Feature Extraction**

Ekstraksi fitur pada teks merupakan proses yang sangat penting terutama untuk teks mining dan information retrieval [15]. Pada penilitian ini akan digunakan model skip-gram dari Word2Vec. Word2Vec diusulkan oleh google, sebuah neural network yang dapat memproses data teks [14]. Skip-Gram adalah sebuah metode untuk mencari konteks dari sebuah kata dalam sebuah kalimat dari kata-kata yang bertetangga. Cara Skip-Gram mendapatkan konteks adalah dengan cara membuat “fake task” yaitu neural network dengan input sebuah kata, 1 hidden layer dan output sebuah kata dengan dimensi yang sama. Fake tasknya adalah “diberikan input sebuah kata ditengah kalimat, lihat kata tetangganya dan pilih secara acak. Neural network ini akan menghasilkan probabilitas untuk setiap kata bertetanggaan dengan kata yang lain. Setiap kata pada neural network akan memiliki hidden layersnya masing2, dan inilah yang akan digunakan sebagai word embeddingnya.

Pada penelitian ini akan digunakan 2 model word embedding, yaitu general word embedding dan domain word embedding. General word embedding akan dibuat menggunakan dataset dari Wikipedia data dump [16] yang berisi teks dari aritkel wikipedia. Dataset ini biasa digunakan karena dapat mencakup banyak kata baku. Tetapi karena data ulasan restoran banyak berisi kata-kata yang tidak baku, maka akan dibuat juga domain word embedding yang akan dibuat dari ulasan review restoran dari website pergikuliner [9]. Kedua word embedding ini akan dibuat dengan model skip-gram dari Word2Vec dengan window size 10 dengan negative sample 5, dimensi untuk general word embedding adalah 200 dan 100 untuk domain word embedding, dengan epoch 15 untuk keduanya, general dan domain.

**Domain Specific Word Embedding**

Untuk membuat Domain Specific Word Embedding, akan akan dibuat data tambahan yaitu data ulasan dari website pergikuliner [9] dengan restoran yang berada di Jakarta dan Bandung. Dari daerah bandung, dipilih 5 daerah akan menjadi kata kunci pencarian pada website, yaitu Raiu, Pasir Kaliki, Pasteur, Dipatiukur, dan Dago. Dari 5 daerah bandung tersebut, terkumpul sekitar 5000 ulasan restoran. Karena jumlahnya masih sedikit, ditambah lagi menggunakan ulasan yang berada di 5 daerah Jakarta yaitu Jakarta Selatan, Utara, Barat, Timur, dan Pusat. Dari kedua daerah tersebut digabung, didapatkan total 85.294 ulasan restoran. Data tersebut akan digabungkan dengan data training pada penelitian sebelumnya [], sehingga terdapat 88.830 data ulasan restoran yang akan digunakan untuk membuat word embedding.

Word Embedding akan dibuat menggunakan model Skip-Gram Word2Vec model dikarenakan model tersebut dapat bekerja lebih baik dengan data yang kecil dan dapat merepresentasikan kata dengan baik meskipun dengan frekuensi yang sedikit [23]. Pembuatan Word Embedding menggunakan dataset ini dapat disebut Domain Specific Word Embedding dikarenakan domain pada data yang spesifik, yaitu ulasan pada restoran di indonesia, sementara word embedding seringnya menggunakan data Wikipedia yang sangat luas domainnya. Dengan menggunakan domain yang spesifik ini, word embedding dapat mengenali kata-kata yang sering digunakan pada review dan jarang dilihat di wikipedia, contohnya kata “kece” (yang memiliki makna keren/bagus), “kane” (enak dibalik), “Instagramable” (mendeskripsikan suasana), dll. Sehingga model akan mendapatkan konteks yang lebih banyak terhadap suatu kalimat dibandingkan ketika menggunakan word embedding biasa yang mungkin tidak mengenali kata-kata tersebut.

**Classification**

Klasifikasi dlakukan dengan algoritma LSTM dan Bi-LSTM. Kedua algoritma tersebut mendapatkan hasil terbaik berdasarkan penelitian sebelumnya [6] dan akan digunakan dalam beberapa skenario pengujian untuk menentukan metode teks preprocessing yang terbaik pada dataset yang digunakan dalam permasalahan ekstraksi aspek. Untuk LSTM akan dibuat dengan 128 hidden layer dengan 0.2 dropout dan 0.2 recurrent dropout, sementara untul Bi-LSTM akan dibuat dengan 128 hidden layer dengan 0.5 dropout dan 0.3 recurrent dropout. Nilai tersebut diambil setelah dilakukan beberapa eksperimen dengan nilai yang berbeda. Berikut arsitektur neural networknya:

Diagram

Description automatically generated

Seperti riset sebelumnya [6], inputnya merupakan n-gram dengan n=3 (tri-gram). Lalu akan diteruskan ke word embedding yang memiliki 2 pilihan percobaan yaitu general atau domain. Selanjutnya akan diteruskan kepada model, pada penilitian ini akan digunakan LSTM dan Bi-LSTM. Average Pooling 1d digunakan untuk pooling dari 3 dimensi kata input menjadi 1. Dan layer terakhir merupakan output dengan softmax dense layer dengan Panjang seesuai dengan jumlah label yaitu 11.

**Evaluasi dan Analisis**

Evaluasi hasil klasifikasi dari algoritma LSTM dan Bi-LSTM akan diukur menggunakan metrik utama yaitu F1-Measure. Perhitungan metrik berikut dapat dilakukan dengan bantuan *confusion matrix*. Skor dari *accuracy, precision, recall* dan *F1-Measure* memiliki rentang nilai 0 hingga 1. Skor dengan nilai mendekati 0 menandakan performa yang buruk, sementara skor dengan nilai mendekati 1 menandakan performa yang baik [13].

|  |  |
| --- | --- |
| *Accuracy =* | (7) |

|  |  |
| --- | --- |
| *Precision =* | (8) |

|  |  |
| --- | --- |
| *Recall =* | (9) |
| *F1-Measure =* | (10) |

Table

Description automatically generated

Gambar 6 *Confusion Matrix [13]*

1. **Evaluasi**

**Hasil Pengujian**

Hasil eksekusi dari masing-masing skenario pengujian dapat dilihat pada tabel. Perhitungan F1-score dibulatkan dengan 2 desimal dan akan dihitung metrics dengan semua label dan tidak menggunakan label O dikarenakan label O adalah mayoritas tag dan sangat mempengaruhi perhitungan F1-score. Algoritma dengan performa terbaik ditandai dengan *shading* berwarna hijau.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Pre-Processing | Word Embedding | Out-of-Vocabularry | Model | F1-Score (Semua Tag) | F1-Score (Tanpa tag-O) |
| Basic | [6] Rachmansyah, General | 17.16% | LSTM | 79.10 | 55.10 |
| Bi-LSTM | 78.90 | 53.00 |
| Domain | 8.49% | LSTM | 75,68 | 45,5 |
| Bi-LSTM | 79,32 | 56,2 |
| Basic + Alay Dictionary (Model 1) | General | 8.11% | LSTM | 75,44 | 44,2 |
| Bi-LSTM | 79,29 | 55,5 |
| Domain | 3.86% | LSTM | 75,34 | 44,4 |
| Bi-LSTM | 79,53 | 56,3 |
| Basic + Alay Dictionary + Stemming (Model 2) | General | 7.54% | LSTM | 75,46 | 45,1 |
| Bi-LSTM | 79,04 | 54,3 |
| Domain | 3.87% | LSTM | 74,85 | 43,5 |
| Bi-LSTM | 78,78 | 52,9 |
| Basic + Alay Dictionary + Rule-Based Processing (Model 3) | General | 8.05% | LSTM | 75,54 | 44,9 |
| Bi-LSTM | 79,28 | 55,3 |
| Domain | 3.62% | LSTM | 75,93 | 45,6 |
| Bi-LSTM | 79,54 | 57,1 |

Berikut merupakah hasil F1-Score tanpa Tag-O untuk model 1 sampai model 3 dengan model Bi-LSTM dan domain word embedding:

|  |  |  |  |
| --- | --- | --- | --- |
| Tag | Model 1 (Bi-LSTM dan Domain Embedding) | Model 2 (Bi-LSTM dan Domain Embedding) | Model 3 (Bi-LSTM dan Domain Embedding) |
| B-FOOD | 0.453 | 0.430 | 0.499 |
| I-FOOD | 0.638 | 0.635 | 0.646 |
| B-MISCELLANEOUS | 0.118 | 0.108 | 0.136 |
| I- MISCELLANEOUS | 0.425 | 0.374 | 0.401 |
| B-SERVICE | 0.549 | 0.459 | 0.444 |
| I- SERVICE | 0.609 | 0.631 | 0.634 |
| B-AMBIENCE | 0.525 | 0.463 | 0.486 |
| I- AMBIENCE | 0.568 | 0.559 | 0.578 |
| B-PRICE | 0.545 | 0.439 | 0.667 |
| I- PRICE | 0.610 | 0.634 | 0.642 |

**Analisis Hasil Pengujian**

Berdasarkan hasil pengujian, pengurangan Out-of-Vocabulary (OOV) sangatlah berpengaruh terhadap performansi model untuk permasalahan aspect extraction, ini didukung oleh riset Anhar, Refany [21] yang membuktikan bahwa menangani OOV dapat meningkatkan performa model dikarenakan model mendapatkan informasi lebih banyak dan detail dibandingkan ketika tidak menangani masalah OOV. Metode-metode pre-processing yang dilakukan dapat mengurangi OOV dengan signifikan dibandingkan dengan penelitian sebelumnya [6]. Model LSTM pada penelitian ini memiliki performa lebih buruk dibandingkan penelitian sebelumnya secara konsisten disetiap percobaan, maka dari itu untuk analisis selanjutnya, hanya akan dilakukan untuk Bi-LSTM dikarenakan hasil yang cukup konsisten ketika dibandingkan dengan penelitian sebelumnya.

Penggunaan domain word embedding memberikan dampak baik terhadap F1-Score dengan rata2 meningkatkan 0.22% pada Semua Tag dan 1,70% tanpa tag-O. Peningkatan ini dikarenakan domain embedding berisi dengan kata-kata yang sering digunakan dalam review makanan dan cenderung tidak baku, sedangkan general embedding yang dibuat dengan dataset Wikipedia cenderung memiliki kata-kata yang baku. Dengan menggunakan word embedding dari kata-kata yang sering digunakan pada review juga membuat OOV pada domain embedding lebih sedikit dibandingkan dengan general embedding, sehingga domain embedding dapat memberikan lebih banyak konteks ke model.

Metode Pre-processing juga sangat mempengaruhi performansi model. Dengan adanya alay dictionary, performansi model meningkat ketika dibandingkan dengan basic preprocessing sebanyak 0.22%, dan lebih baik lagi ketika diberikan rule-based processing yang meningkatkan F1-Score sebanyak 0.33% ketika menggunakan semua tag dan 0.8% ketika tidak menghiraukan tag-O. Sementara ketika menggunakan Stemming, pada penilitian ini performanya mengalami sedikit penurunan, menurut penulis ini dikarenakan stemming dapat mengurangi konteks kalimat, contohnya penggunaan “ditempat” yang biasa digunakan pada review untuk memberi tau kalau penulisnya makan ditempat, sementara kata “tempatnya” biasa diikuti dengan opini penulis mengenai tempat tersebut. Dan ketika stemming dilakukan, maka kedua kata tersebut akan menjadi kata dasarnya yaitu “tempat”.

1. **Kesimpulan**

Pada tugas akhir ini, metode pre-processing basic + alay dictionary + rule-based processing dengan model Bi-LSTM dengan domain word embedding dapat memberikan performa terbaik dengan F1-Score 79.54% dan 57.1% ketika tidak menghiraukan tag-O, dan model ini lebih baik dibandingkan dengan model terbaik pada penelitian sebelumnya [6] dengan F1-Score 79.10 dan 55.10 ketika tidak menghiraukan tag-O. Pada domain review restoran, Domain word embedding dapat memberikan performansi yang lebih baik dibandingkan dengan General word embedding. Untuk Pre-processing, kedua metode Alay Dictionary dan Rule-Based Processing dapat berdampak baik terhadap performansi model dan juga normalisasi harga pada Rule-Based Processing dapat meningkatkan performa model pada memprediksi tag B-PRICE maupun I-PRICE, tetapi pada penelitian ini stemming membuat performansi model menurun. Penurunan jumlah out-of-vocabulary juga berbanding lurus dengan performansi model, lebih sedikit out-of-vocabulary maka model akan memiliki performa yang lebih baik.

**Daftar Pustaka**

1. D. H. Sasmita, A. F. Wicaksono, S. Louvan and M. Adriani, "Unsupervised aspect-based sentiment analysis on Indonesian restaurant reviews," 2017 International Conference on Asian Language Processing (IALP), Singapore, 2017, pp. 383-386.
2. Parikh, Anish, et al. "Motives for reading and articulating user-generated restaurant reviews on Yelp. com." Journal of Hospitality and Tourism Technology (2014).
3. Liu, Bing. "Sentiment analysis and opinion mining." Synthesis lectures on human language technologies 5.1 (2012): 1-167.
4. Schouten, Kim, and Flavius Frasincar. "Survey on aspect-level sentiment analysis." IEEE Transactions on Knowledge and Data Engineering 28.3 (2015): 813-830.
5. Chauhan, Ganpat Singh, et al. "A two-step hybrid unsupervised model with attention mechanism for aspect extraction." Expert Systems with Applications 161 (2020): 113673.
6. Widhianto, Rachmansyah Adhi, and Ade Romadhony. "Aspect Term Extraction Using Deep Learning-Based Approach on Indonesian Restaurant Reviews." 2021 9th International Conference on Information and Communication Technology (ICoICT). IEEE, 2021.
7. Da’u, Aminu, and Naomie Salim. "Aspect extraction on user textual reviews using multi-channel convolutional neural network." PeerJ Computer Science 5 (2019): e191.
8. Tran, Trang Uyen, Ha Thanh Thi Hoang, and Hiep Xuan Huynh. "Aspect extraction with bidirectional GRU and CRF." 2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF). IEEE, 2019.
9. [Online] Available at: <https://pergikuliner.com> [Accessed 15 December 2021]
10. N. Aliyah Salsabila, Y. Ardhito Winatmoko, A. Akbar Septiandri and A. Jamal, "Colloquial Indonesian Lexicon," 2018 International Conference on Asian Language Processing (IALP), 2018, pp. 226-229, doi: 10.1109/IALP.2018.8629151.
11. W Wibowo, Haryo Akbarianto, et al. "Indocollex: A testbed for morphological transformation of indonesian word colloquialism." Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021 (2021).
12. Mohammed
13. Narkhede S. 2018. Understanding Confusion Matrix. [Online]. Available at: <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62> [Accessed 12 December 2021]
14. L. Ma and Y. Zhang, "Using Word2Vec to process big text data," 2015 IEEE International Conference on Big Data (Big Data), 2015, pp. 2895-2897, doi: 10.1109/BigData.2015.7364114.
15. Liang, Hong, et al. "Text feature extraction based on deep learning: a review." EURASIP journal on wireless communications and networking 2017.1 (2017): 1-12.
16. [Online] Available at: https://dumps.wikimedia.org [Accessed 20 December 2021]
17. L. Akhtyamova, P. Martínez, K. Verspoor and J. Cardiff, "Testing Contextualized Word Embeddings to Improve NER in Spanish Clinical Case Narratives," in IEEE Access, vol. 8, pp. 164717-164726, 2020, doi: 10.1109/ACCESS.2020.3018688.Pedregosa
18. R. Perera, T. Malepathirana, Y. Abeysinghe, Y. Albar and U. Thayasivam, "Amalgamation of General and Domain Specific Word Embeddings for Improved Hierarchical Aspect Aggregation," 2019 IEEE 13th International Conference on Semantic Computing (ICSC), 2019, pp. 55-62, doi: 10.1109/ICOSC.2019.8665518.
19. Hanafiah, Novita, et al. "Text normalization algorithm on twitter in complaint category." Procedia computer science 116 (2017): 20-26.
20. Septiandri, Ali Akbar, and Okiriza Wibisono. "Detecting spam comments on Indonesia’s Instagram posts." Journal of Physics: Conference Series. Vol. 801. No. 1. IOP Publishing, 2017.
21. ANHAR, REFANY. MODIFIKASI ARSITEKTUR DEEP LEARNING MODEL PADA NER BAHASA INDONESIA DALAM MENANGANI MASALAH OUT-OF-VOCABULARY. Diss. Universitas Gadjah Mada, 2020.
22. Church, Kenneth Ward. "Word2Vec." Natural Language Engineering 23.1 (2017): 155-162.
23. A Mikolov, Tomas, et al. "Efficient estimation of word representations in vector space." arXiv preprint arXiv:1301.3781(2013).
24. Smagulova, Kamilya, and Alex Pappachen James. "A survey on LSTM memristive neural network architectures and applications." The European Physical Journal Special Topics 228.10 (2019): 2313-2324.
25. Graves, Alex, Santiago Fernández, and Jürgen Schmidhuber. "Bidirectional LSTM networks for improved phoneme classification and recognition." International conference on artificial neural networks. Springer, Berlin, Heidelberg, 2005.
26. Zhang, Yue, Qi Liu, and Linfeng Song. "Sentence-state lstm for text representation." arXiv preprint arXiv:1805.02474 (2018).
27. Ghaeini, Reza, et al. "Dr-bilstm: Dependent reading bidirectional lstm for natural language inference." arXiv preprint arXiv:1802.05577 (2018).