WORD EMBEDDING-BERT UNTUK EKSTRAKSI FITUR PADA ANALISIS SENTIMEN

SKRIPSI

disusun oleh

Aditiya Rahman

17.11.1639

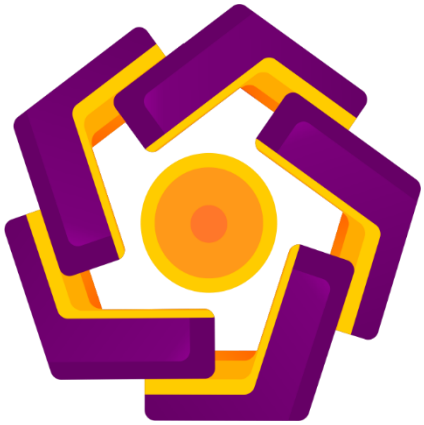
PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2021

WORD EMBEDDING-BERT UNTUK EKSTRAKSI FITUR PADA ANALISIS SENTIMEN

SKRIPSI

Untuk memenuhi sebagian persyaratan
Mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Aditiya Rahman

17.11.1639

PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021

PERSETUJUAN

SKRIPSI

WORD EMBEDDING-BERT UNTUK EKSTRAKI FITUR PADA ANALISIS SENTIMEN

yang dipersiapkan dan disusun oleh

Aditiya Rahman

17.11.1639

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 5 Maret 2021

Dosen Pembimbing,

Mardhiya Hayaty, S.T., M.Kom.,
NIK : 190302108

PENGESAHAN

SKRIPSI

WORD EMBEDDING-BERT UNTUK EKSTRAKI FITUR PADA ANALISIS SENTIMEN

yang dipersiapkan dan disusun oleh

Aditiya Rahman

17.11.1639

telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Februari 2021

Susunan Dewan Penguji

Nama Penguji	**Tanda Tangan**
Uyock Anggoro Saputro, M.Kom. **NIK : 190302419**	
Arif Dwi Laksito, M.Kom. **NIK : 190302150**	
Mardhiya Hayaty, S.T., M.Kom. **NIK. 190302108**	

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Februari 2021

DEKAN FAKULTAS ILMU KOMPUTER

Krisnawati, S.Si, M.T.
NIK. 190302038

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Pangkalpinang, 5 Maret 2021

Aditiya Rahman

NIM. 17.11.1639

MOTTO

“Dreaming as you will live forever. Live as you will die today.”

- James Dean

“You have limited time. Don't waste it by living someone else's life.”

- Steve Jobs

“Pleasure in a work makes perfection at the results achieved.”

- Aristoteles

“The only way to do great work is to love what you do. If you haven’t found it yet, keep looking. Don’t settle.”

- Steve Jobs

PERSEMBAHAN

Alhamdulillah dengan kerja keras serta doa, skripsi ini dapat diselesaikan dengan baik. Segala puji dan syukur bagi Allah SWT yang tiada henti memberikan keberkahan. Dengan ini saya mempersembahkan skripsi ini kepada semua pihak yang terlibat secara langsung atau tidak langsung, yaitu untuk :

1. Kedua orang tua dan kakak – kakak saya yang selalu mendoakan, selalu menyemangati dan memberikan motivasi tiada henti kepada saya.
2. Dosen pembimbing saya Ibu Mardhiya Hayaty, S.T., M.Kom., yang telah membimbing saya dari awal sampai akhir pembuatan skripsi.
3. Dosen-dosen Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu selama masa kuliah.
4. Teman-teman satu kloter penelitian yang memberikan dukungan, bantuan dan motivasi, serta menemani saya untuk menyelesaikan skripsi ini.
5. Teman-teman kelas 17-IF-11 yang telah menemani dan selalu memberikan semangat untuk menyelesaikan skripsi ini.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat serta hidayah-Nya dan shawalat serta salam juga tidak lupa penulis panjatkan kepada junjungan kita Nabi Muhammad SAW yang telah memberikan teladan mulia dalam menuntun umatnya sehingga penulis dapat menyelesaikan skripsi ini.

Skripsi yang berjudul "***WORD EMBEDDING* BERT UNTUK EKSTRAKSI FITUR PADA ANALISIS SENTIMEN**" ini disusun sebagai salah satu syarat utama untuk menyelesaikan program sarjana pada Universitas AMIKOM Yogyakarta. Penyelesaian skripsi ini juga tidak lepas dari bantuan berbagai pihak, karena itu pada kesempatan ini penulis ingin menyampaikan rasa hormat dan terima kasih kepada :

1. Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Bapak Sudarmawan, M.T. selaku ketua Program Studi Informatika Universitas AMIKOM Yogyakarta.
4. Ibu Mardhiya Hayaty, S.T., M.Kom. selaku dosen pembimbing yang selalu bijaksana memberikan bimbingan, nasehat serta waktunya selama penulisan skripsi ini.
5. Bapak Uyock Anggoro Saputro, M.Kom. dan Bapak Arif Dwi Laksito, M.Kom. selaku dosen penguji. Terimakasih atas saran yang diberikan

selama pengujian untuk memperbaiki penelitian menjadi lebih baik lagi.

Penulis menyadari skripsi ini masih ada kekurangan. Maka, penulis menerima kritik dan saran yang membangun serta teguran dari semua pihak. Penulis menerima dengan lapang dada untuk kesempurnaan karya selanjutnya. Semoga skripsi yang sederhana ini bisa bermanfaat, khususnya bagi penulis dan pembaca yang budiman pada umumya. Apabila terdapat kesahalan semoga Allah SWT melimpahkan magfirah-Nya. *Aamiin*.

Pangkalpinang, 5 Maret 2021

Aditiya Rahman

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

INTISARI

Perkembangan teknologi informasi saat ini telah mempermudah kehidupan manusia dalam berbagai bidang, salah satunya di bidang pariwisata yaitu pemesanan hotel. Kita bisa mendapatkan banyak informasi mengenai hotel tersebut mulai dari fasilitas yang diberikan hingga ulasan oleh pengunjung yang sudah pernah menggunakan hotel tersebut. Analisis sentimen merupakan salah satu cabang dari *Natural Language Processing* (NLP) yang dapat membantu dalam menentukan kualitas layanan hotel yang ditawarkan dari ulasan yang telah diberikan pengguna. Penelitian ini menggunakan data ulasan hotel untuk melakukan analisis sentimen yang didapatkan dari situs Traveloka.

Penelitian ini memanfaatkan sebuah metode *deep learning* yaitu *Bidirectional Encoder Representation from Transformer* atau BERT sebagai metode *word embedding* untuk mempresentasikan kata menjadi vektor. Klasifikasi pada penelitian ini dilakukan dengan menambahkan layer *linear regression* pada layer paling atas di BERT.

Dari percobaan yang dilakukan dengan pembagian data *training* sebanyak 70%, data eval sebanyak 10%, dan data tes sebanyak 20% dari total 10.000 data, dapat dilihat bahwa metode BERT memiliki hasil akurasi sebesar 97%, sedangkan metode non-BERT menghasilkan akurasi sebesar 93%. Hasil yang diberikan oleh metode BERT lebih tinggi daripada hasil yang diberikan oleh metode non-BERT. Sehingga bisa disimpulkan bahwa ada pengaruh dari penggunaan metode BERT terhadap akurasi yang dihasilkan.

Kata-kunci: Analisis Sentimen, *Word Embedding*, *Deep Learning*, BERT.

ABSTRACT

The current development of information technology has made human life easier in various fields, one of which is tourism, namely hotel reservations. We can get a lot of information about the hotel from the facilities provided to reviews by visitors who have used the hotel. Sentiment analysis is a branch of Natural Language Processing (NLP) that can help determine the quality of hotel services offered from reviews that users have provided. This study uses hotel review data to analyze sentiments obtained from the Traveloka website.

This study utilizes a deep learning method, namely Bidirectional Encoder Representation from Transformer or BERT as a word embedding method to present words as vectors. The classification in this study was carried out by adding a linear regression layer to the top layer of BERT.

From experiments conducted with the distribution of training data as much as 70%, evaluation data as much as 10%, and test data as much as 20% of the total 10,000 data, it can be seen that the BERT method has an accuracy of 97%, while the non-BERT method produces an accuracy of 93 %. The results given by the BERT method are higher than the results given by the non-BERT methods. So it can be concluded that there is an effect of using the BERT method on the resulting accuracy.

Keywords: *Sentiment Analysis, Word Embedding, Deep Learning, BERT.*