

## ANALISIS SENTIMEN PADA ULASAN HOTEL DI BANDUNG UTARA DENGAN METODE DEEP BELIEF NETWORK

### SENTIMENT ANALYSIS ON HOTEL REVIEW USING DEEP BELIEF NETWORK

Iftah Aziz, Budhi Irawan, Ashri Dinimaharawati

Prodi S1 Teknik Komputer, Fakultas Teknik Elektro, Universitas Telkom  
iftahaziz@student.telkomuniversity.ac.id, budhiirawan@telkomuniversity.co.id,  
ashridini@telkomuniversity.ac.id

---

#### Abstrak

Hotel merupakan sarana tempat tinggal sementara bagi orang yang sedang melakukan perjalanan. Khususnya di Indonesia perkembangan hotel sangat pesat, ini disebabkan oleh meningkatnya pariwisata di Indonesia. Di era internet seperti saat ini dapat memudahkan wisatawan untuk mencari hotel tempat untuk menginap sementara. Wisatawan biasanya menceritakan pengalaman selama berada di hotel dengan menuliskan ulasan di internet. Oleh karena itu banyak sekali ulasan yang ditemukan di internet. Dampaknya ulasan tersebut akan dijadikan acuan oleh pihak hotel untuk memperbaiki dan melakukan evaluasi pada hotelnya, dan wisatawan pun dapat melihat sejauh mana pelayanan hotel tersebut dari ulasan yang ada di internet. Dengan banyaknya ulasan yang beredar di internet, wisatawan harus meluangkan waktu lebih untuk membacanya, dan apakah ulasan tersebut positif atau negatif. Pada penelitian ini dibuat sebuah sistem analisis sentimen pada ulasan hotel dengan metode *Deep Belief Network* dan pembobotan menggunakan *Term Frequency – Inverse Document Frequency* (TF – IDF) yang dapat mengklasifikasikan ulasan hotel kedalam ulasan positif dan negatif. Hasil dari penelitian menggunakan *Deep Belief Network* menghasilkan nilai pengujian tertinggi yang didapat dalam penelitian ini dengan pengujian partisi data dengan menggunakan *learning rate* sebesar 0.001 dan menghasilkan *accuracy* sebesar 93%.

**Kata kunci :** Hotel, analisis sentimen, *Deep Belief Network*, *Term Frequency – Inverse Document Frequency* (TF – IDF).

---

#### Abstract

The hotel is a means of temporary accommodation for people who are traveling. Particularly in Indonesia, the development of hotels is very rapid, this is caused by increased tourism in Indonesia. In the era of the internet as it is today, it can make it easier for tourists to find hotels where to stay overnight. Tourists usually tell their experiences while in the hotel by writing reviews on the internet. Therefore a lot of reviews found on the internet. The impact of these reviews will be used as a reference by the hotel to improve and evaluate the hotel, and tourists can see the extent of the hotel services from the reviews available on the internet. With so many reviews circulating on the internet, tourists have to spend more time reading it, and whether the review is positive or negative. In this study, a sentiment analysis system was made in hotel reviews using the *Deep Belief Network* method and weighting using *Term Frequency - Inverse Document Frequency* (TF - IDF) which can classify hotel reviews into positive and negative reviews. The results of the study using the *Deep Belief Network* produced the highest test value obtained in this study by testing the partition data using a learning rate of 0.001 and producing an accuracy of 93%.

**Keywords:** (Hotel, Sentiment Analysis, *Deep Neural Network*, *Term Frequency – Inverse Document Frequency* (TF – IDF)).

---

#### 1. Pendahuluan

Seiring dengan perkembangan teknologi yang sangat pesat, menjadikan perkembangan pariwisata khususnya di Indonesia pun sejalan dan terus berkembang setiap tahun, hal ini terlihat dari performanya yang semakin meningkat. Pada 2016, devisa pariwisata mencapai 13,5 miliar *dollar* AS per tahun. Selanjutnya pada 2017, sumbangan devisa dari sektor pariwisata melesat menjadi sekitar 16,8 miliar *dollar* AS [1]. Dan di 2019 devisa di sektor pariwisata mencapai 17,6 miliar *dollar* AS [2]. Hal ini menunjukkan sektor pariwisata menjadi sektor yang sangat menjanjikan. Di dalam pariwisata, hotel merupakan bagian yang tidak terpisahkan, hotel ialah Sarana tempat tinggal untuk wisatawan umum dengan memberikan pelayanan jasa kamar, menyediakan makanan, minuman dan akomodasi dengan syarat pembayaran (Lawson, 1976). Hotel juga dapat menjadi salah satu jenis investasi yang mampu memberikan keuntungan. Dalam laporan terbaru bertajuk *Hotel Investment Outlook*, JLL (Jones Lang Lasalle) memaparkan tren positif investasi perhotelan di Asia Pasifik sudah dimulai sejak tahun 2018. [3].

Namun, perkembangan hotel juga dipengaruhi oleh ulasan. Ulasan adalah suatu tipe informasi produk yang dibuat oleh orang-orang yang pernah menggunakan produk tersebut [2]. Ulasan biasanya berisi pendapat, pandangan atau sentimen yang didasarkan pada pengalaman pribadi. Ulasan tersebut dapat dianalisa sehingga bisa

dihasilkan sebuah keluaran yang dapat bermanfaat, dan salah satu model analisa yang dapat dilakukan adalah analisis sentimen.

Analisis sentimen merupakan bagian dari *opinion mining* [2]. Analisis sentimen adalah bidang ilmu yang mempelajari bagaimana menganalisa opini, sentimen, evaluasi, penilaian, sikap dan emosi dari sebuah entitas yang dapat berupa produk, pelayanan, organisasi, individu, isu-isu, peristiwa, dan topik [2]. Analisis sentimen merupakan riset komputasional dari opini, sentimen, dan emosi yang diekspresikan secara tekstual. Sederhananya, analisis sentimen bertujuan untuk klasifikasi teks dari sebuah data, kalimat, atau konten lainnya menjadi beberapa kelompok.

## 2. Dasar Teori

### 2.1 Klasifikasi pada text

Klasifikasi pada teks telah dipelajari secara luas di berbagai daerah, seperti *data mining*, *database*, pembelajaran mesin dan pengambilan informasi [6]. Klasifikasi pada teks diperlukan untuk pengambilan, analisis, akurasi dan anotasi pada teks. Para peneliti sebelumnya telah mempelajari dan mengembangkan berbagai metode untuk klasifikasi pada teks [7]. Klasifikasi pada teks biasa disebut dengan *Supervised Learning*. *Supervised Learning* adalah salah satu teknik dalam pembelajaran mesin untuk mengelompokkan data dari sebuah data latih yang bertujuan untuk menghasilkan prediksi data yang tak terlihat [6].

### 2.2 TF - IDF

*Term Frequency – Inverse Document Frequency* (TF-IDF) adalah statistik numerik yang memiliki maksud untuk menggambarkan pengaruh dari sebuah kata ataupun kalimat dalam sebuah teks. TF-IDF banyak digunakan untuk pembobotan suatu kata atau kalimat. *Term Frequency* adalah proses dimana kata atau kalimat diukur untuk menentukan seberapa sering kata atau kalimat tersebut muncul pada sebuah dokumen. Sedangkan *Invers Document Frequency* adalah proses untuk mengukur seberapa penting istilah kata tersebut dalam dokumen [14]. Untuk mengukur *term frequency* pada *term*  $t$  di dokumen  $d$  dapat di rumuskan sebagai berikut:

$$Tf(t, d) = c(t, d) \quad (2.1)$$

Dimana:

$Tf$  = *Term Frequency*

$t$  = term (Istilah yang dicari)

$d$  = document ( data keseluruhan)

$c$  = frekuensi  $t$  dalam  $d$

Dimana  $c(t, d)$  adalah jumlah frekuensi  $t$  dalam dokumen  $d$ . Untuk menormalkan persamaan 2.1 terhadap jumlah data dalam dokumen maka :

$$Tf(t, d) = \frac{c(t, d)}{l(d)} \quad (2.2)$$

Pada persamaan 2.2 hasil  $Tf$  adalah pembagian dari  $c(t, d)$  yaitu jumlah frekuensi  $t$  dalam dokumen  $d$  dibagi dengan  $l(d)$  yaitu jumlah keseluruhan data dalam dokumen.

Untuk menghitung *inverse document frequency* dapat dirumuskan sebagai berikut:

$$Idf(t) = 1 + \log \frac{N}{K} \quad (2.3)$$

Dalam persamaan 2.3 untuk menghitung nilai IDF atau *inverse document frequency* yaitu dengan  $\log N$  yaitu jumlah keseluruhan dokumen yang dianalisa, dibagi dengan  $K$  yaitu jumlah dokumen dimana istilah yang dicari itu ditemukan.

Dapat kita lihat bahwa, IDF mengukur seberapa sering istilah  $t$  terjadi pada seluruh dokumen yang kita teliti. Jika istilah  $t$  sangat sering muncul pada seluruh dokumen maka  $t$  tersebut tidak dianggap penting dan tidak dapat menjadi yang mewakili dokumen tersebut. Namun sebaliknya, jika  $t$  tersebut jarang muncul pada seluruh dokumen maka  $t$  tersebut dianggap penting dan dapat menjadi yang mewakili untuk di diteliti (karena nilai dari  $idf$  yang tinggi) [15].

Jika kita menghubungkan rumus  $Tf$  dan  $Idf$  maka dihasilkan :

$$weight(t, d) = tf(t, d) * idf(t) \quad (2.4)$$

Dalam persamaan 2.4 *weight* atau bobot dari  $(t, d)$  dihitung dengan melalui perkalian  $tf$  dan  $idf$ .

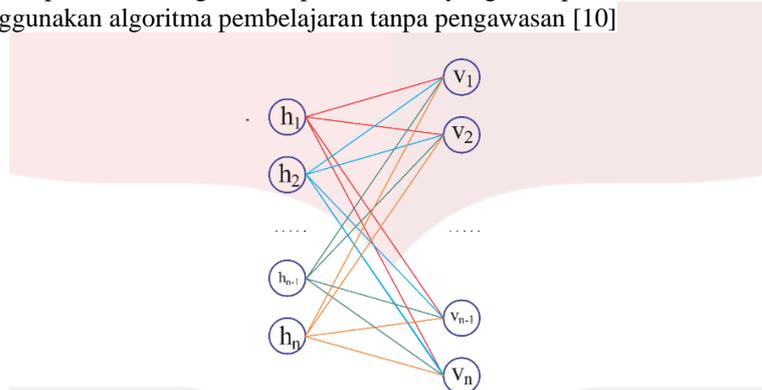
$$\begin{cases} \frac{c(t,d)}{I(d)} * (1 + \log \frac{N}{K}) & \text{Jika } c(t,d) \geq 1 \\ 0 & \text{Lainnya} \end{cases} \quad (2.5)$$

Ket :  $weight(t, d)$  = bobot keseluruhan  $t$  terhadap  $d$

Dalam persamaan 2.5 terdapat penjabaran persamaan 2.4. Persamaan ini menunjukkan Ketika dalam suatu data tidak terdapat suatu *term*, maka data tersebut berbobot nol (0).

**2.3 Restricted Boltzmann Machine**

*Restricted Boltzmann Machine* (RBM) merupakan aturan pembelajaran menggunakan metode *Boltzmann Machine* [10]. RBM merupakan model generatif probabilistik yang mampu secara otomatis mengekstrak fitur input data dengan menggunakan algoritma pembelajaran tanpa pengawasan [10]



Gambar 1 Arsitektur dasar *Restricted Boltzmann Machine*

Pada gambar 1 RBM terdiri dari 2 lapisan, yang pertama adalah lapisan terlihat (*visible layer*) ( $v_1, v_2, \dots, v_m$ ), dan Lapisan tidak terlihat (*hidden layer*) ( $h_1, h_2, \dots, h_n$ ). Baik lapisan terlihat dan tidak terlihat terhubung satu sama lain.

Dari gambar 2.1 jumlah koneksi berkurang dan proses *training* dari *network* ini memiliki banyak persyaratan. Dapat dilihat, jenis jaringan ini memiliki arsitektur yang sangat sederhana, yang merupakan blok bangunan dari *deep belief network*. Fungsi energi dari RBM dapat dilihat di persamaan (2.6).

$$E(v, h) = - \sum_{ij} w_{ij}v_ih_j - \sum_i a_i v_i - \sum_j b_j h_j \quad (2.6)$$

Dimana  $a_i$  adalah bias dari *visible layer*  $i$  yang terlihat dalam keadaan  $v_i$ ,  $b_j$  adalah bias dari *hidden layer*  $j$  yang dalam keadaan  $h_j$  dan  $w_{ij}$  adalah bobot koneksi antara *neuron*  $i$  dan  $j$ .

Probabilitas seluruh sistem dapat disajikan menggunakan keadaan *neuron* di *hidden layer* -  $h$  serta keadaan *neuron* di *visible layer* -  $v$ .

$$\begin{aligned} p(v, h) &= \frac{1}{Z} e^{-E(v,h)} \\ Z &= \sum_{v,h} e^{-E(v,h)} \end{aligned} \quad (2.7)$$

Dari persamaan 2.7  $Z$  merangkum semua pasangan yang memungkinkan dari *hidden layer* dan *visible layer*, dan disebut fungsi partisi. Ini adalah spesialisasi dari *Boltzmann Distribution* ke arsitektur RBM. Persamaan ini memberi peluang untuk menghitung probabilitas yang diaktifkan setiap *neuron*. Diberikan *vector input*  $v$ , probabilitas untuk *hidden layer* yang diaktifkan dapat dilihat di persamaan 2.8.

$$p(h_j = 1|v) = \frac{1}{1 + e^{-(b_j + \sum_i v_i w_{ij})}} = \sigma (b_j + \sum_i v_i w_{ij}) \quad (2.8)$$

Dari persamaan 2.8  $\sigma$  adalah fungsi *sigmoid*. Sejalan dengan ini, probabilitas *visible layer*  $i$  dapat dihitung menggunakan persamaan 2.9.

$$p(h_j = 1|h) = \frac{1}{1 + e^{-(a_j + \sum_i v_i w_{ij})}} = \sigma (a_j + \sum_i v_i w_{ij}) \quad (2.9)$$

$v$  : lapisan terlihat

$h$  : lapisan tidak terlihat

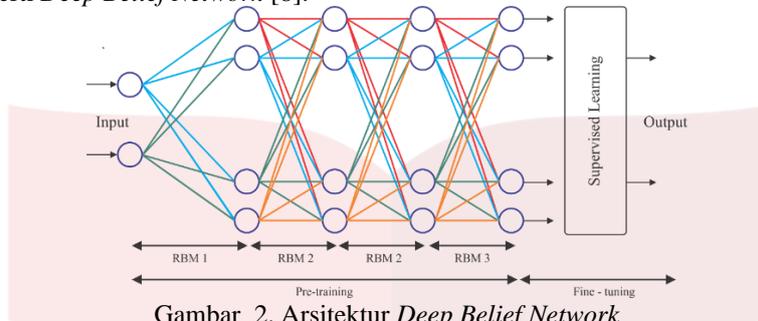
$W$ : bobot dari lapisan terlihat ke lapisan tidak terlihat

$b$ : bias pada lapisan terlihat

$c$ : bias pada lapisan tidak terlihat

**2.4 Deep Belief Network**

Deep Belief Network (DBN) adalah suatu pengembangan dari Deep Learning yang merupakan tumpukan atau stack dari beberapa algoritma atau metode yang bertujuan feature extraction yang memanfaatkan seluruh resource seoptimal mungkin. Deep Learning mencakup algoritma unsupervised dan supervised learning sehingga dapat memanfaatkan data yang berlabel maupun tidak berlabel. Pendekatan yang sering digunakan untuk mengimplementasikan Deep Learning adalah graphical methods atau multi layer representation atau multi layer graphical model seperti Deep Belief Network [8].



Gambar 2. Arsitektur Deep Belief Network

Pada gambar 2, dapat dilihat bahwa arsitektur Deep Belief Network dapat dilihat sebagai komposisi dari sampel yang tidak berlabel seperti Restricted Boltzmann Machines (RBM) [9]. Dimana setiap Hidden Layer akan menjadi visible layer untuk proses selanjutnya. [10]. Dan setelah itu dilakukan proses supervised learning agar dapat mengoptimalkan hasil dari proses klasifikasi.

**2.5 Confusion Matrix**

Evaluasi dilakukan untuk menguji hasil dan klasifikasi yang didapatkan, dengan mengukur nilai performansi dari sudah dibuat. Parameter pengujian yang dipakai untuk evaluasi yaitu akurasi yang perhitungannya berdasarkan tabel confusion matrix. Tabel 2.1 menampilkan sebuah tabel confusion matrix untuk pengklasifikasian kedalam dua kelas [16].

Tabel 2. Confusion Matrix

		True Class	
		Positive	Negative
Predicted Class	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

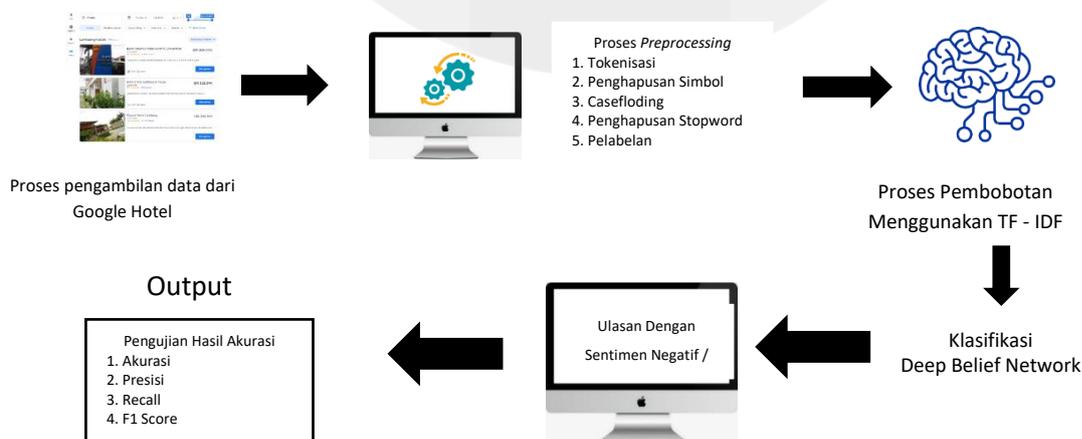
Matriks tersebut mempunyai empat nilai yang dijadikan sebagai acuan dalam perhitungan, yaitu:

1. True Positive (TP), ketika kelas yang diprediksi positif dan faktanya positif.
2. True Negative (TN), ketika kelas yang diprediksi negatif, dan faktanya negatif.
3. False Positive (FP), ketika kelas yang diprediksi positif dan faktanya negatif.
4. False Negative (FN), ketika kelas yang diprediksi negatif dan faktanya positif.

**3. Perancangan**

**3.1. Gambaran Umum Sistem**

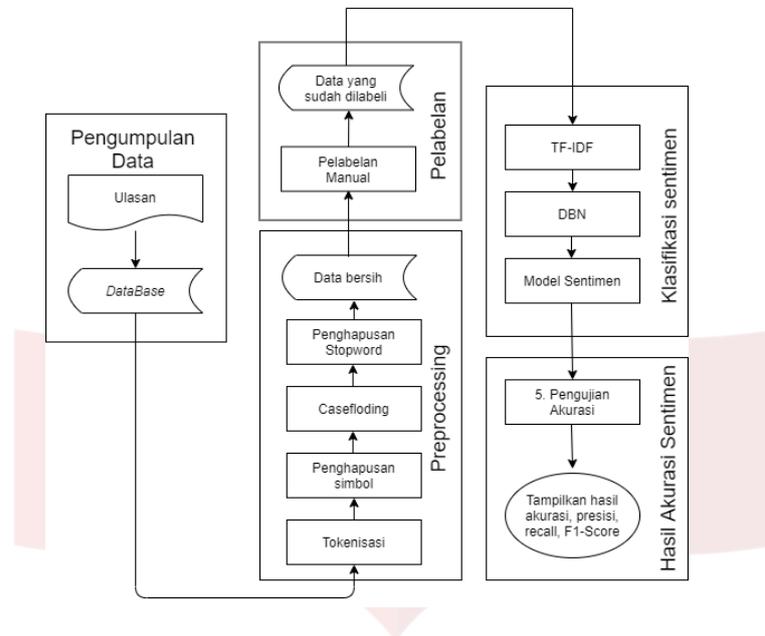
Perancangan proses untuk sistem analisa klasifikasi data ulasan hotel:



Gambar 3. Gambaran Umum Sistem

Pada Gambar 3 menunjukkan proses klasifikasi ulasan hotel berdasarkan sentimen. Data ulasan akan diambil dari situs google *review*. Data yang telah diambil akan langsung masuk ke *database* yang sudah disediakan. Setelah itu data akan diproses menggunakan metode yang sudah dijabarkan. Setelah data diproses, *output* yang akan diperoleh adalah apakah datar tersebut memiliki sentimen positif atau negatif.

Perancangan sistem yang akan dibuat untuk klasifikasi ulasan hotel berdasarkan sentimen adalah sebagai berikut :



Gambar 4 Diagram Alir Perancangan Sistem

Pada Pada Gambar 4 menunjukkan perancangan sistem yang akan dibuat.

### 1. Pengumpulan Data

Pengumpulan data adalah hal yang penting dalam penelitian kali ini. Pengumpulan data dilakukan menggunakan ekstensi google chrome yaitu data miner. Data *miner* dapat mengambil data secara semi otomatis. Cara kerjanya pertama kita cari data yang akan diambil, setelah itu buka aplikasi data miner. Setelah itu arahkan pointer data miner ke data yang akan diambil. Setelah itu cari *parent* atau tipe data yang sama. Setelah itu data dapat diambil secara otomatis.

Pengumpulan data diambil bertahap antara bulan November hingga maret. Pengumpulan data secara bertahap dimaksudkan agar data ulasan akan selalu ter *update*. Setelah data diambil maka data akan disimpan di *database excel*.

### 2. Text Preprocessing

Pada tahap ini data akan dibersihkan, proses yang akan dilewati dari tahap ini adalah proses tokenisasi yaitu pencacahan kalimat yang keluaran nya akan menjadi kata kata yang terpisah. Yang kedua adalah penghapusan simbol yaitu proses menghapus simbol simbol seperti tanda tanya, tanda seru, titik, koma, dan lain sebagainya. Setelah itu data akan melalui proses *casefolding* yaitu membuat data menjadi huruf kecil. Dan yang terakhir data akan diubah menjadi kata dasarnya dalam proses penghapusan *stopword*.

Setelah data melalui proses *preprocessing* data yang sudah bersih akan disimpan Kembali di excel menggunakan extensi xls.

### 3. Pelabelan

Setelah melalui proses *text preprocessing*, maka setelah itu data akan dilabeli secara manual sesuai dengan kelas yang dipersiapkan yaitu kelas ulasan positif dan kelas ulasan negatif. Pelabelan ini dimaksudkan agar nantinya sistem dapat belajar dari data yang sudah dilabeli. Serta nantinya akan digunakan untuk proses pengujian.

### 4. Klasifikasi Sentimen

Pada tahapan ini data akan melewati tahap TF – IDF. Tahap ini digunakan untuk memberikan pembobotan kata. Agar dapat memisahkan kata yang penting dan kata yang tidak penting agar dapat mengoptimalkan proses pada saat klasifikasi.

Setelah pembobotan kata selanjutnya data akan melalui proses utamanya yaitu proses klasifikasi menggunakan metode *Deep Belief Network*. metode ini menggunakan pendekatan *supervised learning*.

Data hasil proses klasifikasi akan masuk ke proses pengujian menggunakan tabel *confusion matrix*. data akan diuji seberapa bagus model klasifikasi yang sudah dilakukan.

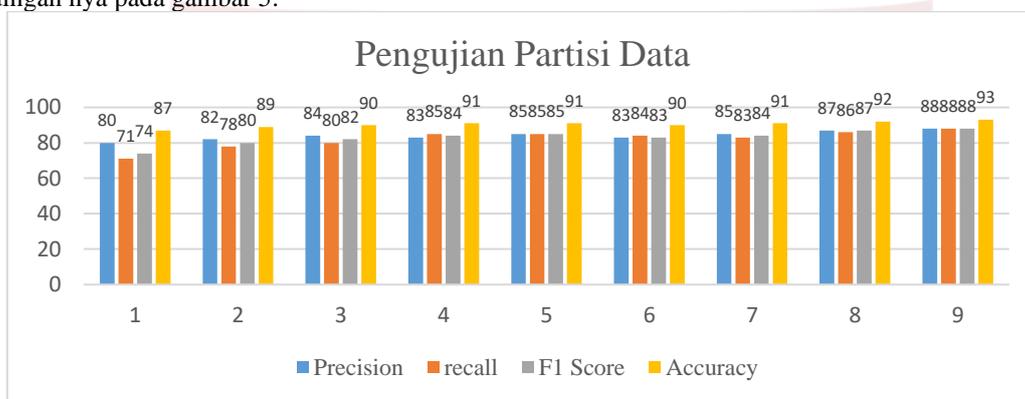
### 4. Pengujian

#### 4.1 Pengujian Partisi data

Tabel 3. Hasil Pengujian Partisi Data

Pengujian Ke-	Data Latih (%)	Data uji (%)	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
1	10	90	80	71	74	87
2	20	80	82	78	80	89
3	30	70	84	80	82	90
4	40	60	83	85	84	91
5	50	50	85	85	85	91
6	60	40	85	84	83	90
7	70	30	85	83	84	91
8	80	20	87	86	87	92
9	90	10	88	88	88	93

Berdasarkan hasil pengujian partisi data pada sistem klasifikasi ulasan hotel, didapatkan bahwa pengujian 9 dengan partisi data 90% data latih dan 10% data uji mendapatkan hasil terbaik dengan nilai accuracy = 93%, Precision = 88%, Recall = 88%, dan F1 Score = 88% dari tabel hasil pengujian tersebut, dapat dilihat grafik perbandingannya pada gambar 5.



Gambar 5. Grafik Pengujian Partisi Data

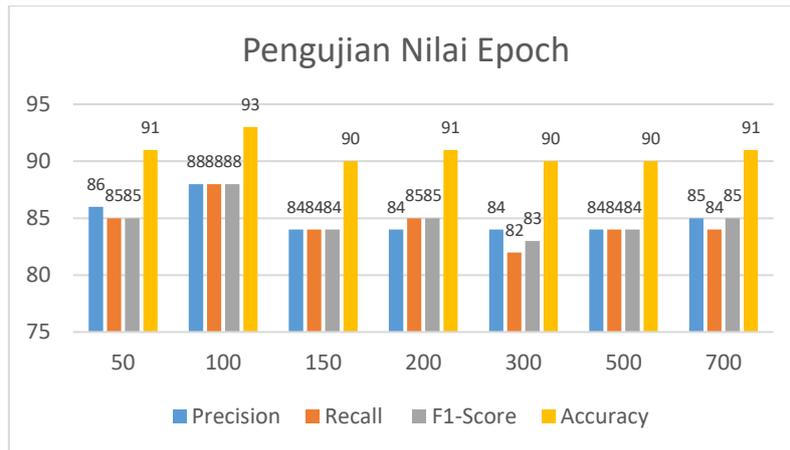
#### 4.2 Pengujian Epoch

Tabel 4. Hasil Pengujian Epoch

Pengujian Ke-	Jumlah Epoch	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
1	50	86	85	85	91
2	100	88	88	88	93
3	150	84	84	84	90
4	200	84	85	85	91
5	300	84	82	83	90
6	500	84	84	84	90
7	700	85	84	85	91

Berdasarkan tabel 4 pada saat pengujian dengan nilai *epoch* 100 menjadi pengujian terbaik dengan nilai akurasi 93%, nilai presisi 88%, nilai *Recall* 88%, dan nilai *F1 Score* 88%. Hal ini menunjukkan bahwa tidak selamanya nilai *epoch* yang tinggi berpengaruh pada performansi. Tidak ada nilai standar untuk *Epoch*, untuk itu pengujian nilai *epoch* berpengaruh pada tingkat performansi sistem.

Pada gambar 4.4 menunjukkan bahwa tidak selamanya *Epoch* yang tinggi berpengaruh pada performansi sistem, dan tidak ada nilai standar untuk *epoch*, menjadikan setiap sistem memiliki nilai optimal *epoch* yang berbeda. Hal ini membuat pengaturan nilai *epoch* yang optimal dapat meningkatkan performansi pada sebuah sistem.



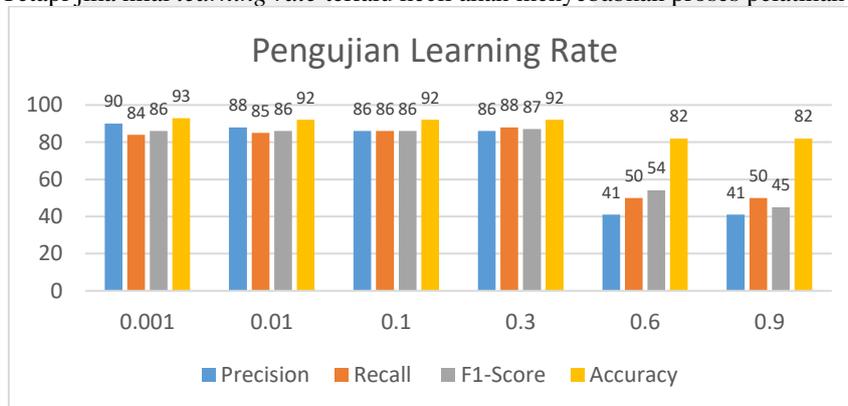
Gambar 6. Grafik Penguujian Epoch

4.3 Penguujian Learning Rate

Tabel 5 Hasil Penguujian Learning Rate

Penguujian Ke-	Learning Rate	Precision (%)	Recall (%)	F1 Score (%)	Accuracy (%)
1	0.1	86	86	86	92
2	0.01	88	85	86	92
3	0.001	90	84	86	93
4	0.3	86	88	87	92
5	0.6	41	50	54	82
6	0.9	41	50	45	82

Berdasarkan tabel 5 pada saat penguujian dengan parameter nilai Learning Rate 0.001 menjadi penguujian terbaik dengan nilai akurasi = 93%, nilai presisi = 90%, nilai Recall = 84%, dan nilai F1 Score = 86%. Hal ini menunjukkan bahwa semakin kecil nilai learning rate maka tingkat performansi dari sistem semakin baik. Hal ini disebabkan karena penggunaan nilai learning rate untuk menentukan seberapa cepat sistem belajar dari data latih yang diperoleh. Tetapi jika nilai learning rate terlalu kecil akan menyebabkan proses pelatihan semakin lama.



Gambar 7. Grafik Penguujian Learning Rate

4.3 Penguujian Response Time

Berdasarkan tiga parameter dalam penguujian sebelumnya, parameter terbaik didapat pada saat learning rate sebesar 0.001, partisi data sebesar 90% data latih dan 10% data uji, dan nilai epoch sebesar 100. Maka dilakukan penguujian respond time guna melihat waktu yang terbaik pada metode Deep belief Network ini. Penguujian dapat dilihat pada tabel di bawah ini:

Tabel 6 Penguujian Response Time

Penguujian Ke	Response Time (detik)
1	96.52
2	95.01
3	96.87
4	95.67
5	97.02
6	95.12
7	96.34

Pengujian Ke	Response Time (detik)
8	94.98
9	97.15
10	96.64
Rata rata (detik)	96.132

Pada tabel 6 dapat dilihat bahwa pengujian *response time* dari dataset ulasan menggunakan metode *Deep Belief Network*. Rata rata setiap pengujian membutuhkan waktu klasifikasi sebesar 96.132 detik atau 1 menit 36 detik. Dari perolehan nilai tersebut dapat disimpulkan bahwa metode *deep belief network* sudah baik dalam melakukan klasifikasi data ulasan hotel, karena pada pengujian ini data yang dipakai mencapai 2004 data ulasan.

## 5. Kesimpulan

Berdasarkan hasil penelitian, pengujian dan analisa yang telah dilakukan pada tugas akhir ini, maka dapat ditarik kesimpulan sebagai berikut:

1. Pada sistem Analisis Sentimen terhadap ulasan hotel menggunakan metode klasifikasi *Deep Belief Network* berbasis web berhasil mengklasifikasikan data ulasan hotel kedalam 2 kelas yaitu positif dan negatif.
2. Berdasarkan Pengujian partisi data, semakin banyak data latih maka *accuracy* klasifikasi yang dihasilkan semakin bagus. Dalam kasus ini partisi data yang terbaik pada proses pengujian yaitu 90% data latih sebanyak 1803 data sedangkan 10% data uji yaitu sebanyak 201 data yang menghasilkan *accuracy* 93%.
3. Nilai *epoch* terbaik yaitu pada pengujian *epoch* dengan nilai 100, dimana memperoleh *accuracy* 93%. Dan nilai *Learning Rate* terbaik didapatkan pada saat pengujian dengan nilai 0.001, dimana memperoleh *accuracy* 93%.

## Daftar Pustaka:

- [1] Cahyu. 2018. "Industri pariwisata Indonesia kian meningkat pesat". *liputan6.com*, 24 September 2018. [Online]. Available: <https://www.liputan6.com/lifestyle/read/3650849/industri-pariwisata-indonesia-kian-meningkat-pesat>. [Accessed 10 September 2019].
- [2] databoks. 2018. "Pendapatan Devisa Indonesia dari Sektor Pariwisata (2009-2019E)". *databoks*, 10 September 2018. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2018/09/10/berapa-pendapatan-devisa-dari-sektor-pariwisata-indonesia>. [Accessed 10 September 2019].
- [3] Putra, Adhitya. "Tahun 2019, Investasi Hotel di Asia Pasifik Akan Meningkat 15%". *Rumahhokie.com*, 13 Maret 2019. [Online]. Available: <https://rumahhokie.com/beritaproperti/tahun-2019-investasi-hotel-di-asia-pasifik-akan-meningkat-15/>. [Accessed 10 September 2019].
- [4] Sulastiyono, Agus. 2010. *Teknik dan Prosedur Divisi Kamar Pada Bidang Hotel*. Bandung: Alfabeta.
- [5] Pemerintah Indonesia. 1986. Surat Keterangan nomor KM.37/PW.304/MPPT-86 BAB 1 pasal 1 ayat (8) tahun 1986 yang mengatur tentang hotel. Lembaran Negara RI tahun 1986, No KM.37. Sekretariat Negara. Jakarta.
- [6] M. Allahyari. E. D. Trippe. and J. B. Gutierrez. "A Brief Survey of Text Mining : Classification, Clustering and Extraction Techniques." 2017.
- [7] Arifin Agus Zainal , Setiono Ari novan (2002). *Klasifikasi dokumen berita kejadian berbahasa Indonesia dengan algoritma single pass clustering*. Proceeding of seminar on intelegent Technology and Its Applications (SITIA), Teknik Elektro, Institut Teknologi Sepuluh.
- [8] K.Kowsari, D.E. Brown. M. Heidaryasafa, K.J. Meimandi. M.S. Gerber. and L. E. Barnes. "HDLTex: Hierarchical Deep Learning for Text Classification," pp.365-371. 2017.
- [9] Yuming Hua, Junhai Guo, and Hua Zhao, —Deep Belief Networks and deep learning, I Proc. 2015 Int. Conf. Intell. Comput. Internet Things, pp. 1–4, 2015 [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7111524>. [Accessed: 6Mar-2019].
- [10] Hinton G (2009). "Deep belief networks". *Scholarpedia*. 4 (5): 5947. doi:10.4249/scholarpedia.5947.
- [11] Suryadi. (2017). *Sentiment Analysis Hotel Review Using Naïve Bayes Classifier Alogorithm And Lexicon Based Approach*. Tesis. Tidak Diterbitkan. Fakultas Matematika Dan Ilmu Pengetahuan Alam. Universitas Gadjah Mada. Yogyakarta.
- [12] Zulfaa, Ira. Winarko, Edi. (2017). *Sentimen Analisis Tweet Berbahasa Indonesia dengan Deep Belief Network*.
- [13] Ida, Bagus, Gede. Made, Sudarma. I, Nyoman, Satya, Kumara, Asti. (2016). *Klasifikasi Teks Bahasa Bali Dengan Metode Supervised Learning Naïve Bayes Classifier*.
- [14] R. Xu, 'POS weighted TF-IDF algorithm and its application for an MOOC search engine', in *2014 International Conference on Audio, Language and Image Processing*, Shanghai, China, Jul. 2014, pp. 868–873, doi: 10.1109/ICALIP.2014.7009919.
- [15] N. A. Setyadi, M. Nasrun, and C. Setianingsih, 'Text Analysis For Hate Speech Detection Using Backpropagation Neural Network', in *2018 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC)*, Bandung, Indonesia, Dec. 2018, pp. 159–165, doi: 10.1109/ICCEREC.2018.8712109.