

VISUALISASI ANALISA SENTIMEN OBJEK

PARIWISATA RAJA AMPAT

SKRIPSI



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

Diajukan Guna Memenuhi Persyaratan Memperoleh

Gelar Sarjana Komputer (S.Kom.)

Samrida Adeldytalia

00000017643

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNIK DAN INFORMATIKA

UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2020

PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi ini adalah karya ilmiah saya sendiri, bukan plagiat dari karya ilmiah yang ditulis oleh orang lain atau lembaga lain, dan semua karya ilmiah orang lain atau lembaga lain yang dirujuk dalam skripsi ini telah disebutkan sumber kutipannya serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan / penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah skripsi yang telah saya tempuh dan status kesarjanaan strata satu yang sudah diterima akan dicabut.

Tangerang, 7 Juli 2020



Samrida Adeldytalia

HALAMAN PENGESAHAN

Skripsi dengan judul

“Visualisasi Analisa Sentimen Objek Pariwisata Raja Ampat”

oleh

Samrida Adeldytalia

telah diujikan pada hari Selasa, 23 Juni 2020,

pukul 13.00 – 14.30 dan dinyatakan lulus

dengan susunan penguji sebagai berikut

Ketua Sidang



Friska Natalia, Ph.D.

Penguji



Suryasari, S.Kom., M.T

Dosen Pembimbing



Melissa Indah Fianty, S.Kom., M.MSI.

Disahkan oleh

Ketua Program Studi Sistem Informasi – UMN



Ririn Ikana Desanti, S.Kom., M.Kom.

VISUALISASI ANALISA SENTIMEN OBJEK PARIWISATA RAJA AMPAT

ABSTRAK

Oleh: Samrida Adeldytalia

Berdasarkan data dari Dinas Pariwisata Kabupaten Raja Ampat, setiap tahunnya Raja Ampat memiliki kenaikan jumlah wisatawan yang cukup signifikan. Sebagai objek wisata, Raja Ampat tentu harus memastikan fasilitas pada lokasi wisata memadai dan mampu menjamin turis lokal maupun internasional saat datang ke Raja Ampat merasa puas. Pada jaman sekarang masyarakat biasanya memilih untuk membagikan pengalaman mereka pada saat berwisata di media sosial sebagai wujud penilaian kepuasan pelanggan. Tentunya akan ada berbagai penilaian dari pengunjung, baik tanggapan negatif maupun positif, salah satunya pada *posting-an* media sosial *Twitter* yang dapat mempengaruhi kesan terhadap objek pariwisata tersebut. Melihat hal ini, dilakukan penelitian untuk melihat penilaian masyarakat terhadap objek wisata Raja Ampat.

Pada penelitian kali ini, akan digunakan data *tweets* dari media sosial *Twitter* dan metodologi klasifikasi *Naïve Bayes* dengan membagi klasifikasi menjadi 3 yaitu positif, negatif, dan netral. Proses analisa menggunakan Bahasa pemrograman R pada *software RStudio*, lalu hasil akhir ditampilkan dalam bentuk *dashboard* pada *software Power BI*.

Penelitian ini menunjukkan bahwa hasil pegujian akurasi algoritma *Naïve Bayes* sebesar 83.93%, dengan sentimen positif memiliki persentase tertinggi dengan 81.90% yaitu sebanyak 19.691 data, kemudian dilanjut oleh sentimen netral 12.14% dengan data sebanyak 2.918 data dan terakhir sentimen negatif 5.96% dengan hanya 1.433 data. Data ditampilkan melalui visualisasi pada *Power BI* untuk memudahkan melihat hasil analisa.

Kata Kunci : analisa sentimen, *Twitter*, pariwisata, raja ampat, *naïve bayes*

SENTIMENT ANALYSIS VISUALIZATION OF RAJA AMPAT TOURISM

ABSTRACT

By: Samrida Adeldytalia

Based on data from the Raja Ampat tourism office, year to year Raja Ampat has a significant increase number of tourists. As a tourist attraction, Raja Ampat should make sure the facilities are able to guarantee the satisfaction of tourists. Nowadays people usually share their experiences when traveling on social media as a form of customer satisfaction assessment. Of course there will be a variety kind of visitors from the assessment, both negative and positive responses. That way could affect reputation of Raja Ampat. Based on this, then raised the to see the public sentiment of Raja Ampat.

The research will use Twitter tweets as data, and the Naïve Bayes classification methodology will be used by dividing classification into 3, positive, negative, and neutral. The analysis process uses R programming language on RStudio, then the final results are presented in a dashboard on the Power BI software.

The results of this research show that the accuracy of Naïve Bayes algorithm is 83.93%, with positive sentiment having the highest percentage with 81.90% of 19,691 data, then followed by the neutral sentiment 12.14% of 2,918 data and negative sentiment of 5.96% with only 1,433 data. Data is presented through visualization on Power BI to make it easier to see the results of the analysis.

Keywords: sentiment analysis, twitter, tourism, Raja Ampat, naïve bayes

KATA PENGANTAR

Puji dan syukur kepada Allah SWT yang selalu memberikan petunjuk dan pencerahan kepada penulis sehingga skripsi yang berjudul “Visualisasi Analisa Sentimen Objek Pariwisata Raja Ampat” dapat selesai tepat pada waktunya. Skripsi ini penulis ajukan kepada Program Strata 1, Program Studi Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi, Universitas Multimedia Nusantara.

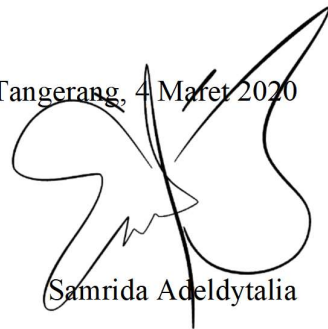
Dengan selesainya proses penulisan skripsi ini, penulis juga ingin menyampaikan ucapan terimakasih kepada:

1. Orangtua saya, Bapak Sukriadi dan Ibu Wilmadiani Rachlina juga tak lupa kakak saya Samara Yarasevika yang telah memberikan semangat dan do’a yang tak henti-hentinya kepada penulis selama penyusunan skripsi ini.
2. Bapak Dr. Ninok Leksono selaku Rektor Universitas Multimedia Nusantara
3. Ibu Ririn Ikana Desanti, S. Kom, M. Kom selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Melissa Indah Fianty, S.Kom., M.MSI. yang telah memberikan bimbingan dan saran-saran kepada penulis selama pengerjaan skripsi,
5. Teman – teman sejawat di perkuliahan yang selalu memberikan *support* dalam berbagai bentuk kepada penulis
6. Teman – teman peserta *training Digital Talent Scholarship 2019* bidang *Big data Analytics* yang telah memberikan saran dan *insight* dengan penulis.

7. Muhammad Hanif Triyana yang membantu dan menjadi tempat diskusi penulis setiap saat.
8. Semua penulis jurnal, buku, dan sumber pemrograman yang telah banyak membantu selama proses penulisan skripsi ini karena secara tidak langsung membuka jendela pengetahuan bagi penulis yang akan berguna bagi masa depan penulis.

Semoga skripsi ini dapat memberikan informasi dan inspirasi yang bermanfaat bagi para pembaca.

Tangerang, 4 Maret 2020

A handwritten signature in black ink, consisting of several fluid, overlapping loops and a long, sweeping stroke that extends upwards and to the right.

Samrida Adeldytalia

DAFTAR ISI

PERNYATAAN	ii
HALAMAN PENGESAHAN	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL	xiv
DAFTAR RUMUS	xv
BAB I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Batasan Masalah	5
1.3. Rumusan Masalah	6
1.4. Tujuan dan Manfaat	6
1.4.1. Tujuan Penelitian.....	6
1.4.2. Manfaat Penelitian.....	7
BAB II LANDASAN TEORI	8
2.1. Raja Ampat	8
2.2. Media Sosial	9

2.3. <i>Big Data</i>	10
2.4. <i>Data Mining</i>	13
2.5. Visualisasi Data.....	14
2.6. Analisa Sentimen	14
2.7. <i>Twitter</i>	16
2.7.1. <i>Twitter API</i>	16
2.8. <i>RStudio</i>	17
2.9. <i>Hashtag</i>	18
2.10. <i>Power BI</i>	19
2.11. <i>Naïve Bayes</i>	19
2.12. Penelitian Terdahulu	21
BAB III METODOLOGI PENELITIAN	24
3.1. Objek Penelitian.....	24
3.2. Alur Penelitian	24
3.2.1. Pengumpulan data	9
3.2.2. <i>Data preprocessing</i>	11
3.2.3. Perancangan model.....	13
3.2.4. Pembuatan <i>Dashboard</i>	13
3.2.5. <i>Problem Solving</i>	14
3.3. Variabel Penelitian.....	16

3.4. Teknik Pengumpulan Data	16
3.5. Teknik Pengambilan Sampel	16
3.6. Teknik Analisis Data	17
BAB IV ANALISIS DAN HASIL PENELITIAN	18
4.1. Data	18
4.2. <i>Data Pre-processing</i>	20
4.2.1. <i>Tokenizing</i>	20
4.2.2. <i>Case folding</i>	22
4.2.3. <i>Spelling Normalization</i>	24
4.2.4. <i>Stopword Removal</i>	25
4.2.5. <i>Stemming</i>	26
4.3. Perancangan Model	28
4.4. Visualisasi	35
4.5. Hasil Diskusi	44
BAB V KESIMPULAN DAN SARAN	45
5.1. Kesimpulan	45
5.2. Saran	47
DAFTAR PUSTAKA	48
LAMPIRAN	54

DAFTAR GAMBAR

Gambar 1.1. Jumlah Wisatawan Raja Ampat 2014-2018.....	2
Gambar 1.2. Data pengguna media sosial di Indonesia per Januari 2020	4
Gambar 2.1. Raja Ampat.....	8
Gambar 2.2. Hierarki tugas analisis sentiment.....	15
Gambar 2.3. Logo <i>Twitter</i>	16
Gambar 2.4. Tahapan mendapatkan akses <i>API Twitter</i>	17
Gambar 3.1. Alur penelitian	25
Gambar 3.2. Tampilan awal setelah masuk <i>Twitter developer</i>	9
Gambar 3.3. Mendaftar pembuatan aplikasi	9
Gambar 3.4. Terdapat token untuk pengambilan data	10
Gambar 3.5. <i>Script</i> untuk mengambil <i>tweets</i>	10
Gambar 3.6. proses <i>pre-processing</i> data.....	11
Gambar 4.1. contoh <i>tweets</i> yang mengandung <i>keyword #RajaAmpat</i>	18
Gambar 4.2. <i>Flowchart</i> seleksi data pada <i>Microsoft Excel</i>	19
Gambar 4.3. <i>Script Input</i> data dan menghapus spasi.....	20
Gambar 4.4. <i>Script</i> proses <i>tokenizing</i>	21
Gambar 4.5. <i>Script</i> proses <i>Case folding</i>	23
Gambar 4.6. <i>Script</i> proses <i>Spelling Normalization</i>	24

Gambar 4.7. <i>Script</i> proses <i>Stopword Removal</i>	25
Gambar 4.8. <i>load library</i> yang dibutuhkan.....	27
Gambar 4.9. fungsi penampung kamus katadasaR.....	27
Gambar 4.10. <i>Script</i> dalam tahap <i>labelling</i>	29
Gambar 4.11. Perbandingan hasil klasifikasi	31
Gambar 4.12. Salah satu <i>Script</i> pemisahan data.....	32
Gambar 4.13. <i>Script</i> model <i>Naïve Bayes</i>	33
Gambar 4.14. <i>Script</i> penghitungan tingkat akurasi	33
Gambar 4.15. Hasil tabel <i>confusion matrix</i>	34
Gambar 4.16. Hasil tingkat akurasi <i>Naïve Bayes</i>	35
Gambar 4.17. <i>Input</i> data dengan jenis data csv.....	36
Gambar 4.18. Hasil <i>Input</i> data yang akan digunakan	36
Gambar 4.19. Visualisasi 1	37
Gambar 4.20. Data untuk visualisasi sentimen	38
Gambar 4.21. Visualisasi 2	38
Gambar 4.22. Data untuk visualisasi <i>slicer</i>	39
Gambar 4.23. Visualisasi 3	39
Gambar 4.24. Data untuk visualisasi jumlah <i>tweets</i>	39
Gambar 4.25. Visualisasi 4	40
Gambar 4.26. Data untuk visualisasi rata-rata <i>tweets</i>	40

Gambar 4.27. Rumus <i>tweets</i> perhari.....	40
Gambar 4.28. Visualisasi 5	41
Gambar 4.29. Data untuk visualisasi jumlah <i>likes</i>	41
Gambar 4.30. Visualisasi 6	41
Gambar 4.31. Data untuk visualisasi rata-rata <i>likes</i>	41
Gambar 4.32. Rumus <i>likes</i> per <i>tweets</i>	42
Gambar 4.33. Visualisasi 7	42
Gambar 4.34. Data untuk visualisasi grafik <i>tweets</i>	43
Gambar 4.35. Tampilan <i>Dashboard</i>	43

DAFTAR TABEL

Tabel 2.1. Tabel Penelitian Terdahulu.....	21
Tabel 3.1. Perbandingan <i>Naïve Bayes</i> dengan K-Nearest Neighbor	15
Tabel 4.1. Hasil <i>Tokenizing</i>	21
Tabel 4.2. Hasil proses <i>Case folding</i>	23
Tabel 4.3. Hasil <i>Spelling Normalization</i>	25
Tabel 4.4. Hasil <i>Stopword Removal</i>	26
Tabel 4.5. Hasil proses <i>Stemming</i>	28
Tabel 4.6. Hasil <i>labelling</i>	30
Tabel 4.7. Tabel perbedaan sebelum dan sesudah penelitian.....	44

DAFTAR RUMUS

Rumus 2.1. Rumus perhitungan <i>Naïve Bayes</i>	20
---	----