

SKRIPSI

ANALISIS DATA REVIEW RAMEN



Antonius Yoga Krisanto

NPM: 2017730082

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
2022**

UNDERGRADUATE THESIS

RAMEN REVIEW DATA ANALYSIS



Antonius Yoga Krisanto

NPM: 2017730082

**DEPARTMENT OF INFORMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
2022**

LEMBAR PENGESAHAN

ANALISIS DATA REVIEW RAMEN

Antonius Yoga Krisanto

NPM: 2017730082

Bandung, 6 Januari 2022

Menyetujui,

Pembimbing

Digitally signed
by Veronica Sri
Moertini

Dr. Veronica Sri Moertini, Ir., M.T.

Ketua Tim Penguji
Digitally signed
by Luciana
Abednego

Luciana Abednego, M.T

Anggota Tim Penguji
Digitally signed
by Rosa de Lima
E. Padmowati

Rosa De Lima, M.Kom

Mengetahui,

Ketua Program Studi
Digitally signed
by Mariskha Tri
Adithia

Mariskha Tri Adithia, P.D.Eng

PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

ANALISIS DATA REVIEW RAMEN

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
Tanggal 6 Januari 2022



Antonius Yoga Krisanto
NPM: 2017730082

ABSTRAK

Ramen merupakan mi kuah khas Jepang yang terdiri dari mi yang dibuat dari gandum, daging, kaldu ikan, dan dibumbui miso atau kecap. *Ramen* sudah menjadi salah satu kuliner asal Jepang yang sukses di berbagai negara. Negara lain pun banyak memodifikasi rasa *ramen* ini agar dapat menyesuaikan dengan selera masyarakat di negaranya. Opini atau pendapat dari pelanggan di berbagai *website* dan *social media* menjadi salah satu "kunci" bagi para pemilik bisnis *ramen* untuk menyajikan *ramen* yang diminati oleh masyarakat dinegaranya.

Dalam praktiknya untuk menemukan jenis rasa *ramen* yang disukai oleh masyarakat diperlukan analisis tentang rasa *ramen* melalui *review* atau ulasan yang diunggah oleh para penikmat kuliner di Indonesia. Oleh karena itu untuk membantu masyarakat yang ingin melihat hasil *review* terhadap *ramen* mereka, akan dilakukan analisis statistik dan analisis sentimen pada data *review ramen* berbahasa Indonesia dengan menggunakan teknologi yang sesuai dan dengan menerapkan tahapan-tahapan *data science*. Analisis statistik digunakan untuk melihat hubungan faktor-faktor yang ada untuk meningkatkan penjualan *ramen*. Sedangkan analisis sentimen merupakan salah satu contoh pengaplikasian dari *Natural Language Processing* (NLP) yang akan digunakan untuk menilai emosi positif atau negatif para penikmat *ramen* terhadap *ramen* yang dibeli. Analisis sentimen ini akan menghasilkan sebuah model klasifikasi yang akan dipilih melalui berbagai macam tahapan evaluasi terhadap algoritma-algoritma yang digunakan. Algoritma yang akan diuji pada pemodelan ini adalah *logistic regression*, *decision tree*, *support vector machine*, *naive bayes*, *random forest*. Dan hasil yang didapatkan menunjukkan bahwa algoritma *random forest* dengan jumlah pohon 90 memiliki nilai akurasi yang terbaik dalam mengklasifikasikan sebuah ulasan *ramen* termasuk ke dalam ulasan positif atau negatif, dengan nilai akurasi sebesar 85.2%. Model klasifikasi tersebut akan digunakan pada perangkat lunak *Ramen's Sentiment Analyzer* yang ditujukan kepada para pemilik usaha *ramen* untuk membantu melihat tingkat kepuasan para pelanggan pada *ramen* yang mereka buat.

Skripsi ini bertujuan untuk menghasilkan sebuah perangkat lunak yang dapat menganalisis ulasan dan memanfaatkan hasil analisis untuk memberikan informasi yang dapat membantu para pemilik usaha *ramen* meningkatkan penjualan *ramen*. Dengan menggunakan perangkat lunak tersebut, pengguna dapat melihat nilai sentimen dari sekumpulan teks ulasan *ramen* berbahasa Indonesia dan memanfaatkan hasil analisis sentimen tersebut untuk usaha *ramennya*.

Kata-kata kunci: *ramen*, ulasan, analisis sentimen, *Ramen's Sentiment Analyzer*

ABSTRACT

Ramen is a typical Japanese noodle soup consisting of noodles made from wheat, meat, fish broth, and seasoned with miso or soy sauce. Ramen has become one of the successful Japanese cuisines in various countries. Other countries also modify the taste of this ramen to suit the tastes of the people in their country. Opinions or opinions from customers on various sites and social media are one of the "keys" for ramen business owners to serve ramen that is of interest to the people in their country.

In practice, to find the type of ramen flavor that is liked by the public, it is necessary to analyze the taste of the ramen through reviews or reviews uploaded by culinary connoisseurs at Indonesia. Therefore, to help people who want to see the results of a review of their pramen, statistical analysis and sentiment analysis will be carried out on the Indonesian-language ramen review data using appropriate technology and by applying the stages of scientific data. Statistical analysis used to see the relationship of the existing factors to increase sales of ramen. Meanwhile, sentiment analysis is an example of the application of Natural Language Processing (NLP) which will be used to assess the positive or negative emotions of ramen lovers towards the purchased ramen. This sentiment analysis will produce a classification model that will be selected through various evaluations of the algorithms used. The algorithms that will be tested in this modeling are logistic regression, decision trees, vector support machines, naive bayes, random forests. And the results obtained show that the random forest algorithm with 90 trees has the best accuracy value in classifying a ramen review into positive or negative reviews. The classification model will be used in Ramen's Sentiment Analyzer software which is aimed at ramen business owners to help see the level of customer satisfaction with the ramen they make.

This thesis aims to produce a software that can analyze reviews and utilize the results of the analysis to provide information that can help ramen business owners increase ramen sales. By using the software, users can see the sentiment value of a set of Indonesian-language ramen review texts and take advantage of the results of the sentiment analysis for their ramen business.

Keywords: *ramen, review, sentiment analysis, Ramen's Sentiment Analyzer*

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yesus Kristus karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan penulisan skripsi ini yang berjudul *Analisis Data Review Ramen*. Penulis menyadari bahwa penyelesaian penyusunan skripsi ini juga tidak terlepas dari bantuan berbagai pihak. Penulis ingin berterima kasih kepada :

- Tuhan Yesus Kristus yang telah membimbing penulis dan memberikan dukungan dan bantuan kepada penulis dalam proses penulisan skripsi.
- Kepada ayah saya Benny Januar Pribady dan ibu saya Elisabeth Kristin Setyaningsih dan keluarga besar saya yang selalu memberikan dukungan, doa dan semangat selama saya kuliah;
- Kepada Ibu Dr. Veronica Sri Moertini selaku pembimbing. Terima kasih atas ilmu, pengalaman, semangat, dan bimbingan yang telah diberikan;
- Kepada Ibu Luciana Abednego, M.T. dan Ibu Rosa De Lima. selaku penguji yang memberikan masukan dan saran, sehingga skripsi ini dapat dikembangkan menjadi lebih baik;
- Adik dari penulis Leonardus Bayu Krisanto yang selalu memberikan semangat dan menghibur dalam proses pengerjaan skripsi;
- Kepada para teman-teman mahasiswa bimbingan Fariz, Indra, Fritz, dan Yohan yang sudah berjuang dan saling memberikan semangat hingga lulus bersama;
- Kepada Jason Christian teman seperjuangan di Teknik Informatika yang selalu memberikan semangat dan mendukung saya dalam pengerjaan skripsi dan kuliah;
- Kepada teman-teman mahasiswa angkatan 2017 sampai 2020. Terima kasih atas seluruh momen kebersamaannya.
- Kepada mereka yang tidak bisa disebutkan satu-persatu yang telah membantu dan mendukung penulis selama penulisan skripsi.

Sebagai penutup, penulis berterima kasih kepada pembaca dokumen skripsi ini. Penulis berharap semoga penelitian yang dilakukan dapat bermanfaat bagi kemajuan ilmu pengetahuan, dan dapat menjadi referensi yang baik bagi penelitian di masa depan. Penulis memohon maaf jika terdapat kesalahan penulisan, sistematika, atau metode yang dipaparkan pada skripsi ini. Penulis sangat terbuka untuk menerima masukan dan saran tambahan dari pembaca.

Bandung, Januari 2022

Penulis

DAFTAR ISI

KATA PENGANTAR	xv
DAFTAR ISI	xvii
DAFTAR GAMBAR	xix
DAFTAR TABEL	xxi
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Batasan Masalah	3
1.5 Metodologi	3
1.6 Sistematika Pembahasan	4
2 LANDASAN TEORI	5
2.1 <i>Ramen</i>	5
2.1.1 Asal usul <i>Ramen</i>	5
2.1.2 Jenis-jenis <i>Ramen</i>	6
2.1.3 <i>Ramen</i> di Indonesia	7
2.2 <i>Natural Language Processing (NLP)</i>	8
2.2.1 Tahapan <i>PreProcessing</i> Data Teks	9
2.2.2 <i>Vector Space Model(VSM)</i>	10
2.2.3 Analisis Sentimen	11
2.3 Teknologi <i>Web Scrapping</i>	12
2.4 Visualisasi Data	13
2.5 Teknik Evaluasi Model	16
2.6 Algoritma Klasifikasi	18
2.6.1 <i>Logistic Regression</i>	18
2.6.2 <i>Decision Tree</i>	18
2.6.3 <i>Naive Bayes</i>	20
2.6.4 <i>Support Vector Machine</i>	20
2.6.5 <i>Random Forest</i>	20
3 PENGUMPULAN, EKSPLORASI DATA DAN TEKNOLOGI	23
3.1 Rancangan Solusi Masalah	23
3.2 Dekripsi Dataset	23
3.3 Persiapan Data	25
3.4 Ekplorasi Data	29
3.4.1 Eksplorasi Data 1	29
3.4.2 Hasil Eksplorasi Data 1	31
3.4.3 Eksplorasi Data 2	32

3.5	Hasil Eksplorasi Data 2	42
3.6	Penyiapan Data untuk Analisis Sentimen	42
3.6.1	Deskripsi Dataset	42
3.6.2	Persiapan Data	43
4	PEMBUATAN DAN EVALUASI MODEL	45
4.1	Tahapan Pemodelan	45
4.2	Evaluasi Model	47
4.2.1	Evaluasi Model Tahap Pertama	47
4.2.2	Evaluasi Model Tahap Kedua	52
4.2.3	Evaluasi Model Tahap Ketiga	54
4.3	Pengujian Model Klasifikasi	56
5	PEMANFAATAN MODEL	59
5.1	Perancangan dan Pembuatan Aplikasi	59
5.2	Hasil Aplikasi <i>Ramen's Sentiment Analyzer</i>	63
5.3	Pengujian Aplikasi <i>Ramen's Sentiment Analyzer</i>	66
6	KESIMPULAN DAN SARAN	71
6.1	Kesimpulan	71
6.2	Saran	72
	DAFTAR REFERENSI	73
	A KODE PROGRAM	75

DAFTAR GAMBAR

1.1	<i>Ramen</i> kuliner khas asal Jepang (sumber: https://www.google.com/)	1
1.2	Restoran yang akan dianalisis (sumber: https://www.google.com/)	2
2.1	Komponen <i>NLP</i> (sumber : https://www.tutorialspoint.com/)	9
2.2	Teknologi <i>web scrapping</i> (sumber: parsehub.com dan scrapehero.com)	12
2.3	Hasil <i>scrapping</i> dalam format <i>excel</i> (sumber: parsehub.com)	12
2.4	Tampilan Halaman Utama <i>ScrapeHero</i> (sumber: scrapehero.com)	13
2.5	Contoh Visualisasi Bar Chart	13
2.6	Fungsi Visualisasi Data	13
2.7	Bar chart untuk membandingkan kategori	14
2.8	<i>Pie chart</i> untuk evaluasi hierarki	15
2.9	<i>Line chart</i> untuk menunjukkan perubahan	15
2.10	<i>Scatter plot</i> untuk melihat distribusi data	16
2.11	<i>Confussion Matrix</i>	17
2.12	Gambaran pengambilan keputusan pada <i>Decision Tree</i> (sumber: https://glints.com/)	19
2.13	Cara kerja <i>Decision Tree</i>	19
3.1	Contoh data <i>ramen</i> yang didapat	24
3.2	Dataframe setelah penyiapan tahap 1 dan 2	25
3.3	Tabel Keterangan Data Korpus	26
3.4	Contoh isi Korpus	26
3.5	Hasil eksekusi fungsi <i>Words</i>	27
3.6	Hasil <i>output trans1</i> dan <i>trans2</i>	28
3.7	Pengujian <i>PreProcess</i>	29
3.8	Top 10 topik <i>Yagami Ramen House</i> pada bulan Februari 2021	32
3.9	Top 10 topik <i>Yagami Ramen House</i> pada bulan Maret 2021	33
3.10	Top 10 topik <i>Yagami Ramen House</i> pada bulan April 2021	33
3.11	Ulasan tentang harga <i>ramen</i> <i>Yagami Ramen House</i>	34
3.12	Ulasan tentang penyajian <i>ramen</i> <i>Yagami Ramen House</i>	34
3.13	Top 10 pasangan kata <i>review</i> <i>Yagami Ramen House</i> bulan Februari	35
3.14	Top 10 pasangan kata <i>review</i> <i>Yagami Ramen House</i> bulan Maret	35
3.15	Top 10 pasangan kata <i>review</i> <i>Yagami Ramen House</i> bulan April	35
3.16	Top 10 topik <i>Sai Ramen</i> pada bulan Februari 2021	36
3.17	Top 10 topik <i>Sai Ramen</i> pada bulan Maret 2021	36
3.18	Top 10 topik <i>Sai Ramen</i> pada bulan April 2021	36
3.19	<i>Street Bar</i> ala Jepang <i>Sai Ramen</i>	37
3.20	Ulasan tentang suasana kedai <i>Sai Ramen</i>	37
3.21	Top 10 pasangan kata <i>review</i> <i>Sai Ramen</i> bulan Februari	38
3.22	Top 10 pasangan kata <i>review</i> <i>Sai Ramen</i> bulan Maret	38
3.23	Top 10 pasangan kata <i>review</i> <i>Sai Ramen</i> bulan April	38
3.24	Top 10 topik <i>Nobu Ramen and Sushi</i> pada bulan Februari 2021	39
3.25	Top 10 topik <i>Nobu Ramen and Sushi</i> pada bulan Maret 2021	39
3.26	Top 10 topik <i>Nobu Ramen and Sushi</i> pada bulan April 2021	40

3.27	Top 10 pasangan kata <i>review</i> Nobu Ramen and Sushi bulan Februari	40
3.28	Top 10 pasangan kata <i>review</i> Nobu Ramen and Sushi bulan Maret	41
3.29	Top 10 pasangan kata <i>review</i> Nobu Ramen and Sushi bulan April	41
3.30	Ulasan tentang harga <i>ramen</i> Nobu Ramen and Sushi	42
3.31	Contoh isi dataset untuk analisis sentimen	43
3.32	Format dataset setelah tahap penyiapan data	43
3.33	Sebaran jumlah <i>rating</i> data untuk Analisis Sentimen	44
3.34	Perbandingan jumlah <i>review</i> dengan sentimen positif dan negatif	44
4.1	Hasil <i>Sparse Matrix</i> pada data simetris	46
4.2	Fitur-fitur yang didapat	47
4.3	Perbandingan jumlah <i>review</i> dengan sentimen positif dan negatif	52
4.4	Perbandingan jumlah <i>review</i> dengan sentimen positif dan negatif	55
4.5	<i>Source kode program RandomForestEngine.ipnyb</i>	56
5.1	Implementasi kode untuk <i>single review analysis</i>	60
5.2	Implementasi kode untuk <i>multiple review analysis</i>	60
5.3	Tampilan awal aplikasi <i>Ramen's Sentiment Analyzer</i>	63
5.4	Tampilan <i>single review analysis 1</i>	64
5.5	Tampilan <i>single review analysis 2</i>	65
5.6	Tahap pemilihan file untuk <i>multiple review analysis</i>	65
5.7	Tampilan <i>multiple review analysis</i>	66
5.8	Data Sai Ramen Januari 2022	67
5.9	Pengujian dengan data Sai Ramen Januari 2022	67
5.10	Pengujian dengan data Yagami Ramen House Januari 2022	68
5.11	Pengujian dengan data Nobu Ramen and Sushi Januari 2022	68

DAFTAR TABEL

3.1	Teknik <i>scrapping</i> yang digunakan untuk setiap URL	24
3.2	Atribut-atribut pada dataset <i>ramen</i>	24
3.3	Tabel Cara Kerja Delete	27
3.4	Tabel Frekuensi Kemunculan Kata Yagami Ramen House	29
3.5	Tabel 11 kuah ramen Yagami Ramen House	30
3.6	Tabel Frekuensi Kemunculan Kata Sai Ramen	30
3.7	Tabel Frekuensi Kemunculan Kata Nobu Ramen and Sushi	31
3.8	Atribut-atribut pada dataset <i>ramen</i> untuk Analisis Sentimen	43
4.1	Contoh <i>rating review</i> 3 dengan sentimen negatif	45
4.2	Bag of Words Data Dummy	47
4.3	Hasil Akurasi Model <i>Logistic Regression</i>	48
4.4	Hasil Akurasi Model <i>Decision Tree</i>	48
4.5	Hasil Akurasi Model <i>Naive Bayes</i>	49
4.6	Hasil Akurasi Model <i>Support Vector Machine</i>	49
4.7	Hasil Akurasi Model Random Forest	50
4.8	Hasil Akurasi Algoritma <i>Random Forest</i> jumlah pohon 90	50
4.9	Hasil Akurasi Algoritma <i>Random Forest</i> jumlah pohon 40	51
4.10	Hasil Akurasi Algoritma <i>Random Forest</i> jumlah pohon 80	51
4.11	Hasil Evaluasi Algoritma <i>Random Forest</i> jumlah pohon 40	53
4.12	Hasil Evaluasi Algoritma <i>Support Vector Machine</i>	53
4.13	Hasil Evaluasi Algoritma <i>Random Forest</i> jumlah pohon 90	54
4.14	Rata-rata keseluruhan <i>metrics</i> model	54
4.15	Hasil Evaluasi Algoritma <i>Random Forest</i> jumlah pohon 90	55
4.16	Tabel pengujian model klasifikasi	57

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan bisnis kuliner dari waktu ke waktu cenderung terus meningkat dan semakin beragam. Salah satu bisnis kuliner yang sedang diminati masyarakat pada saat ini adalah *ramen*. *Ramen* merupakan mi kuah khas Jepang yang terdiri dari mi yang terbuat dari gandum, daging, kaldu ikan, dan dibumbui *miso* atau kecap. *Ramen* sudah menjadi salah satu kuliner asal Jepang yang sukses di berbagai negara. Negara - negara di dunia pun banyak memodifikasi *ramen* ini agar dapat menyesuaikan dengan selera masyarakat di negaranya masing-masing. Opini atau pendapat dari pelanggan di berbagai *website* dan sosial media menjadi salah satu "kunci" bagi para pemilik bisnis *ramen* untuk menyajikan *ramen* yang diminati oleh masyarakat di daerahnya.



Gambar 1.1: *Ramen* kuliner khas asal Jepang (sumber:<https://www.google.com/>)

Di Indonesia, *ramen* sudah menjadi salah satu bisnis makanan yang "menjamur" di berbagai daerah. Pada dasarnya setiap daerah di Indonesia umumnya memiliki rasa dan keanekaragaman masakannya masing-masing yang khas. Salah satu contohnya adalah provinsi Jawa Barat, dimana setiap kota dan kabupaten menjual jajanan Sunda dengan rasanya khas dan digemari oleh masyarakat di provinsi Jawa Barat seperti batagor, cilok, cireng, basreng dan lain-lain. Dari contoh tersebut dapat dikatakan bahwa umumnya sebuah kelompok di masyarakat memiliki masakan khas andalan masing-masing yang dari segi rasanya memiliki kecenderungan tertentu, misalnya masakan pedas, manis, ataupun asin. Hal inilah yang membuat rasa *ramen* yang ada di Indonesia memiliki rasa yang bervariasi dengan keunikannya masing-masing.

Oleh karena itu untuk membantu para pemilik bisnis *ramen* untuk membuka bisnis dan meningkatkan penjualan *ramen* khususnya di Bandung, akan dilakukan analisis pada data ulasan *ramen* berbahasa Indonesia yang didapatkan menggunakan teknologi *web scrapping*. *Web Scrapping* adalah sebuah proses pengambilan data dari *website* yang dilakukan baik secara otomatis ataupun manual. Teknologi-teknologi yang digunakan untuk melakukan *scrapping* adalah ParseHub dan ScrapeHero. Kedua teknologi tersebut digunakan untuk mendapatkan data dari berbagai macam

website seperti *Google Review* dan *PergiKuliner.com*. Dari hasil *web scrapping* ini akan didapatkan berbagai ulasan dari tiga restoran *ramen* di Bandung yang dipilih dikarenakan memiliki ciri dan ke khas-an nya masing-masing pada menu *ramen*-nya. Ketiga restoran itu adalah Sai Ramen, Yagami Ramen, dan Nobu Ramen. Sai ramen memiliki ciri khas menu *ramen* yang dikatakan oleh sebagian besar orang mirip dengan *ramen* asli dari Jepang, Yagami ramen dengan menu *ramen* yang lebih kekinian salah satunya adalah menu *tomyam ramen* dan *black pepper ramen*, dan terakhir Nobu Ramen dengan ciri khas semua menu *ramen*-nya yang terbuat dari kaldu ayam dan *topping* ayam.



Gambar 1.2: Restoran yang akan dianalisis (sumber:<https://www.google.com/>)

Analisis yang akan dilakukan adalah analisis statistik dan analisis sentimen. Analisis statistik digunakan untuk melihat hubungan berbagai faktor yang ada untuk meningkatkan penjualan *ramen*. Pada skripsi ini, analisis diawali dengan mencari topik yang sering dibicarakan dari data ulasan yang didapat. Untuk mencari topik yang dibicarakan pada data ulasan yang didapat, digunakan konsep *Natural Language Processing (NLP)* dimana terdapat tahapan *text pre processing* yang digunakan untuk mencari topik pada data ulasan *ramen*. *NLP* adalah sub-bidang dalam ilmu komputer khususnya *Artificial Intelligence* yang memungkinkan komputer dapat memahami dan memproses bahasa alami manusia. Penerapan konsep ilmu *NLP* ini akan sangat membantu dalam melakukan analisis statistik terhadap data ulasan *ramen*. Berikutnya adalah analisis sentimen, analisis sentimen merupakan salah satu contoh pengaplikasian dari *NLP* yang dapat membantu menganalisis sejumlah besar ulasan tentang suatu produk. Pada ulasan *ramen* yang didapat, terdapat makna dan emosi yang disampaikan secara tersirat. Bagi komputer makna dan emosi tersebut dapat dipahami dengan menggunakan analisis sentimen. Pemahaman makna dan emosi dari data ulasan itu penting untuk mengetahui apakah emosi seorang pelanggan kepada *ramen* yang dijual berupa emosi positif, negatif atau netral. Algoritma yang akan diuji dalam pemodelan ini adalah *logistic regression*, *decision tree*, *support vector machine*, *naive bayes*, *random forest*. Dari kelima algoritma tersebut akan dicari algoritma mana yang memiliki nilai akurasi yang terbaik dalam mengklasifikasikan sebuah ulasan *ramen* termasuk ke dalam ulasan positif atau negatif. Untuk mendapatkan model klasifikasi yang terbaik, akan dilakukan evaluasi terhadap model klasifikasi yang didapat. Evaluasi yang dilakukan akan memanfaatkan konsep *confussion matrix* dimana dengan menggunakan konsep tersebut dapat dilihat nilai *presision*, *recall*, dan *f1-score* dari model klasifikasi yang telah dibuat. Model klasifikasi terbaik yang didapat akan digunakan pada perangkat lunak *Ramen's Sentiment Analyzer* yang ditujukan kepada para pemilik usaha *ramen* untuk membantu melihat tingkat kepuasan para pelanggan pada *ramen* yang mereka buat.

Skripsi ini bertujuan untuk menghasilkan sebuah perangkat lunak yang dapat menganalisis ulasan dan memanfaatkan hasil analisis untuk memberikan informasi yang dapat membantu para pemilik usaha *ramen* meningkatkan penjualan *ramen*. Dengan menggunakan perangkat lunak tersebut, pengguna dapat melihat nilai sentimen dari sekumpulan teks ulasan *ramen* berbahasa Indonesia dan memanfaatkan hasil analisis sentimen tersebut untuk usaha *ramennya*.

1.2 Rumusan Masalah

Dari deskripsi masalah yang telah dipaparkan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara mengumpulkan data *review ramen*?
2. Bagaimana tahapan untuk menyiapkan data *review ramen* agar dapat dianalisis?
3. Bagaimana cara melakukan analisis data *review ramen*?
4. Bagaimana cara menginterpretasi dan mengevaluasi hasil analisis data *review ramen*?
5. Bagaimana cara membangun perangkat lunak untuk analisis data *review ramen*

1.3 Tujuan

Dari rumusan masalah yang telah dipaparkan, tujuan dari penelitian ini adalah sebagai berikut:

1. Menggunakan teknologi *web scrapper* untuk mengumpulkan data *review ramen*.
2. Memanfaatkan teknik *text pre-processing* pada *Natural Language Processing* untuk menyiapkan data *review ramen* agar dapat dianalisis.
3. Membuat model klasifikasi dan melakukan analisis sentimen terhadap data *review ramen*.
4. Menginterpretasi dan mengevaluasi data *review ramen* yang telah dianalisis.
5. Membangun perangkat lunak untuk menganalisis data *review ramen* dan memanfaatkan hasil analisisnya.

1.4 Batasan Masalah

Batasan dan asumsi yang digunakan dalam penelitian ini adalah:

1. Data yang digunakan adalah restoran *ramen* Yagami Ramen House, Sai Ramen, Nobu and Sushi Ramen.
2. Data yang digunakan untuk eksplorasi dan pembuatan model klasifikasi berada pada rentang 1 Februari 2021 sampai dengan April 2021.
3. Model analisis sentimen digunakan untuk memprediksi nilai sentimen positif dan negatif.

1.5 Metodologi

Metodologi pada penelitian ini adalah sebagai berikut:

1. Studi Literatur
Pencarian referensi dan sumber-sumber lainnya yang dapat digunakan sebagai acuan dalam pembahasan mengenai *ramen* dan *web scraping*.
2. Pengumpulan Data
Melakukan pengumpulan data *review ramen* dengan menggunakan teknik-teknik *web scraping* yang sudah dipelajari.
3. Eksplorasi dan Persiapan Data
Melakukan tahapan eksplorasi pada data *review ramen* untuk melihat topik yang banyak dibicarakan pada data *review ramen* yang dikumpulkan dan melakukan tahapan persiapan data agar data *review ramen* dapat segera diproses dan dianalisis.
4. Pemodelan dan Evaluasi Data
Melakukan pemodelan klasifikasi data *review ramen* analisis sentimen dan melakukan evaluasi terhadap hasil analisis data *review ramen*.
5. Perancangan dan Pembuatan Program
Perancangan dan pembuatan program untuk melakukan analisis sentimen terhadap data *review ramen* yang menjadi masukan. Program dapat menerima masukan sebuah *review* atau sekumpulan *review* dalam format *.csv*.

6. Penulisan Dokumen

Pembuatan dokumen skripsi tentang penelitian yang dilakukan.

1.6 Sistematika Pembahasan

Sistematika penulisan pada skripsi ini adalah sebagai berikut:

1. Bab 1: Pendahuluan untuk mendefinisikan masalah yang akan dibahas, rumusan masalah, batasan dan asumsi, tujuan penelitian dan metodologi penyelesaian masalah.
2. Bab 2: Dasar-dasar teori yang digunakan dalam penelitian ini. Pembahasan mengenai *ramen*, *Natural Language Processing* dan *web scraping*.
3. Bab 3: Pengumpulan, Eksplorasi Data *review ramen* dan teknologi yang digunakan pada penelitian ini. Pada bab ini juga dipaparkan tahapan yang dilakukan ketika melakukan penyiapan data.
4. Bab 4: Pembuatan dan Evaluasi model dengan menggunakan konsep *confussion matrix* yang dilakukan pada data *review ramen*. Pada bab ini juga akan dipaparkan pengujian yang dilakukan terhadap model klasifikasi yang digunakan untuk analisis sentimen.
5. Bab 5: Pemanfaatan model dimana diimplementasikan sebuah prototipe aplikasi *Ramen's Sentiment Analyzer* yang digunakan untuk mengklasifikasikan nilai sentimen *review ramen*.
6. Bab 6: Kesimpulan dan saran