

**ANALISIS SENTIMEN ULASAN PENGGUNA  
APLIKASI ZOOM MENGGUNAKAN METODE *NAÏVE BAYES***

**SKRIPSI**

**OLEH:  
NUR JIHAN QOTHRUNNADA  
NIM. 16610007**



**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

**ANALISIS SENTIMEN ULASAN PENGGUNA  
APLIKASI ZOOM MENGGUNAKAN METODE *NAÏVE BAYES***

**SKRIPSI**

**Diajukan Kepada  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
untuk Memenuhi Salah Satu Persyaratan  
dalam Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh  
Nur Jihan Qothrunnada  
NIM. 16610007**

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

**ANALISIS SENTIMEN ULASAN PENGGUNA  
APLIKASI ZOOM MENGGUNAKAN METODE NAÏVE BAYES**

**SKRIPSI**

**Oleh  
Nur Jihan Qothrunnada  
NIM. 16610007**

**Telah Disetujui Untuk Diuji  
Malang, 23 Juni 2023**

**Dosen Pembimbing I,**



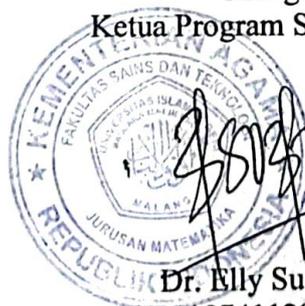
**Hisyam Fahmi, M.Kom.  
NIP. 19890727 201903 1 018**

**Dosen Pembimbing II,**



**Erna Herawati, M.Pd  
NIDT. 19760723 20180201 2 222**

**Mengetahui,  
Ketua Program Studi Matematika**



**Dr. Elly Susanti, M.Sc.  
NIP. 19741129 200012 2 005**

# ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI ZOOM MENGGUNAKAN METODE NAÏVE BAYES

## SKRIPSI

Oleh  
**Nur Jihan Qothrunnada**  
NIM. 16610007

Telah Dipertahankan di Depan Dewan Penguji  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)  
Tanggal 30 Juni 2023

Ketua Penguji : Dr. Elly Susanti, M.Sc.

Anggota Penguji 1 : Muhammad Khudzaifah, M.Si

Anggota Penguji 2 : Hisyam Fahmi, M.Kom

Anggota Penguji 3 : Erna Herawati, M.Pd

"   
.....  
  
.....  
  
.....  
  
.....

Mengetahui,  
Ketua Program Studi Matematika

  
  
Dr. Elly Susanti, M.Sc.  
NIP. 19741129 200012 2 005

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Nur Jihan Qothrunnada  
NIM : 16610007  
Program Studi : Matematika  
Fakultas : Sains dan Teknologi  
Judul Skripsi : Analisis Sentimen Ulasan Pengguna Aplikasi *Zoom*  
Menggunakan Metode *Naïve Bayes*

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 30 Juni 2023

Yang membuat pernyataan,



Nur Jihan Qothrunnada  
NIM. 16610007

## **MOTO**

*"Ketakutan adalah penjara bernama kegagalan. Taklukan rasa takut karena sukses adalah hak pemberani."*

*(Jefri Al Buchori)*

## **PERSEMBAHAN**

Skripsi ini penulis persembahkan untuk:

Ayah tercinta Sahid, ibu tercinta Munifah, kakak-kakak terkasih Mufrotul Faidah dan Misbahul Munir, serta adik tersayang Muhammad Kholil yang senantiasa mendukung, menopang, dan memberikan doa, dukungan, serta motivasi terbaik bagi penulis.

## KATA PENGANTAR

*Assalamu'alaikum Warahmatullahi Wabarakatuh*

Segala puji bagi Allah Swt. yang telah melimpahkan karunia serta rahmat-Nya sehingga peneliti berhasil menyelesaikan penulisan skripsi dengan judul “Analisis Sentimen Ulasan Pengguna Aplikasi *Zoom* Menggunakan Metode *Naïve Bayes*”. Shalawat serta salam semoga terus tercurahkan kepada Nabi Muhammad saw. yang semoga kelak kita semua senantiasa diberi keberkahan dan syafaatnya.

Ucapan syukur dan terima kasih peneliti persembahkan kepada berbagai pihak yang telah memberikan bimbingan, bantuan, serta arahan kepada penulis. Untuk itu, ucapan terima kasih yang sebesar-besarnya penulis sampaikan kepada:

1. Bapak Prof. Dr. H. M. Zainuddin, MA. selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Ibu Dr. Sri Harini, M.Si. selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Ibu Dr. Elly Susanti, M.Sc. selaku Ketua Program Studi Matematika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Bapak Hisyam Fahmi, M.Kom., selaku dosen pembimbing I yang telah memberikan waktu, bimbingan, serta arahan selama perkuliahan hingga dapat terselesaikannya skripsi.
5. Ibu Erna Herawati, M.Pd., selaku dosen pembimbing II serta dosen wali yang telah memberikan bimbingan, nasihat, ilmu, serta arahan kepada penulis.
6. Bapak Prof. Dr. H. Turmudi, M.Si., Ph.D, selaku dosen wali yang senantiasa memberi penulis semangat dan motivasi selama proses menimba ilmu.

7. Seluruh Dosen, Staff Administrasi Jurusan matematika, Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang yang telah memberikan ilmu serta pengalamannya kepada penulis.
8. Orang tua dan seluruh keluarga besar yang selalu senantiasa mendoakan, memberikan dukungan, semangat, serta kasih sayang sehingga penulis dapat menyelesaikan tugas akhir.
9. Kepada semua pihak yang turut membantu dalam menyelesaikan skripsi ini.

Akhirnya penulis berharap semoga skripsi ini dapat bermanfaat bagi penulis dan pembaca.

*Wassalamu'alaikum Warahmatullahi Wabarakatuh*

Malang, 30 Juni 2023

Penulis

## DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PENGAJUAN.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TULISAN.....	v
MOTO.....	vi
PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
ABSTRAK.....	xiii
ABSTRACT.....	xiv
مستخلص البحث.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
1.6 Definisi Istilah.....	4
BAB II KAJIAN TEORI.....	6
2.1 Analisis Sentimen.....	6
2.2 <i>Data Preprocessing</i> .....	7
2.3 <i>Naive Bayes</i> .....	8
2.4 <i>Confusion Matrix</i> .....	9
2.5 Kajian Keislaman.....	11
BAB III METODE PENELITIAN.....	14
3.1 Jenis Penelitian.....	14
3.2 Data dan Sumber Data.....	14
3.3 Tahapan Penelitian.....	15
BAB IV PEMBAHASAN.....	18
4.1 Hasil Tahapan <i>Data Preprocessing</i> .....	18
4.1.1 <i>Data Cleaning</i> dan <i>Case Folding</i> .....	18
4.1.2 <i>Tokenization</i> .....	18
4.1.3 <i>Stopword Removal</i> .....	19
4.1.4 <i>Stemming</i> .....	19
4.2 Hasil Analisis Sentimen Pengguna Aplikasi Zoom.....	20
4.3 Analisis Sentimen dalam Pandangan Islam.....	23
BAB V PENUTUP.....	27
5.1 Kesimpulan.....	27
5.2 Saran.....	27
DAFTAR PUSTAKA.....	28
LAMPIRAN.....	30
RIWAYAT HIDUP.....	33

## DAFTAR TABEL

<b>Tabel 2.1</b> <i>Confusion Matrix</i> .....	9
<b>Tabel 3.2</b> Sampel Data Ulasan Pengguna Zoom .....	14
<b>Tabel 4.1</b> Proses <i>Data Cleaning</i> dan <i>Case Folding</i> .....	18
<b>Tabel 4.2</b> Proses <i>Tokenization</i> .....	18
<b>Tabel 4.3</b> Proses <i>Stopword Removal</i> .....	19
<b>Tabel 4.4</b> Proses <i>Stemming</i> .....	19
<b>Tabel 4.5</b> Hasil Analisis Sentimen Tiap Kelas.....	21
<b>Tabel 4.6</b> Hasil Analisis Sentimen dengan 10-fold cross validation.....	22

## DAFTAR GAMBAR

Gambar 3.1 <i>Flowchart</i> Tahapan Analisis Sentimen .....	17
Gambar 4.1 <i>Confusion Matrix</i> Hasil Analisis Sentimen.....	20

## ABSTRAK

Qothrunnada, Nur Jihan. 2023. **Analisis Sentimen Ulasan Pengguna Aplikasi Zoom Menggunakan Metode *Naive Bayes***. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Hisyam Fahmi, M.Kom. (II) Erna Herawati, M.Pd.

**Kata Kunci:** Analisis Sentimen, *Naive Bayes*, *Data Preprocessing*, Zoom

Penerapan Pembatasan Sosial Berskala Besar (PSBB) dilaksanakan guna mencegah meluasnya penyebaran *Corona Virus Disease-19* (COVID-19). Akibatnya, kegiatan-kegiatan pada bidang pendidikan, ekonomi, sosial maupun budaya harus diselenggarakan secara daring. Salah satu aplikasi *video conference* yang sering digunakan adalah Zoom. Seiring dengan bertambahnya pengguna aplikasi Zoom, saran dan kritik diberikan oleh pengguna melalui ulasan untuk mendapatkan pengalaman yang baik. Analisis sentimen merupakan studi komputasi yang mempelajari opini, sikap, serta perasaan terhadap suatu objek. Hasil dari analisis sentimen kemudian dapat digunakan sebagai bahan pertimbangan dalam mengambil keputusan. Penelitian ini bertujuan untuk mengetahui hasil analisis sentimen pengguna aplikasi Zoom dengan menggunakan metode *naive bayes*. Data ulasan yang diteliti didapatkan dari situs Google Play Store berjumlah 200 data yang terdiri dari 50 ulasan dengan *rating* bintang 1, 2, 4, dan 5. Dataset ulasan harus melalui tahapan *preprocessing* terlebih dahulu dengan menghapus tanda baca, angka, dan kata-kata tidak bermakna yang tidak mempengaruhi tingkat akurasi hasil analisis sentimen. Kemudian, dataset akan diuji dengan mengklasifikasikan sentimen menggunakan metode *Naive Bayes*. Hasil yang diperoleh menunjukkan bahwa model yang dihasilkan bekerja dengan baik dalam menentukan sentimen dari ulasan dengan nilai akurasi sebesar 73,5%.

## ABSTRACT

Qothrunnada, Nur Jihan. 2023. **Sentiment Analysis of Zoom App User Reviews Using Naive Bayes Method**. Thesis. Mathematics Study Program, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisors: (I) Hisyam Fahmi, M.Kom. (II) Erna Herawati, M.Pd.

**Keywords:** Data Preprocessing, Naive Bayes, Sentiment Analysis, Zoom

The implementation of Large-Scale Social Restrictions (PSBB) was carried out to prevent the spread of the Corona Virus Disease-19 (COVID-19). Therefore, activities in the educational, economical, social and cultural sectors must be held online. One of the video conferencing applications that is frequently used is Zoom. Along with the increase in zoom application users, suggestions and criticisms are given by users through reviews to get a good experience. Sentiment analysis is a computational study that studies opinions, attitudes, and feelings towards an object. The results of sentiment analysis can then be used as consideration in making decisions. This study aims to determine the sentiment analysis of zoom application users using the Naïve Bayes method. The reviews used in this study was obtained from Google Play Store site in number of 200 data consisting of 50 reviews with rating score 1, 2, 4, and 5. Firstly, the reviews dataset must go through the preprocessing step by removing punctuations, numbers, and meaningless words that do not affect the accuracy of sentiment analysis. Then, the dataset will be tested by classifying sentiments using the Naive Bayes method. The results showed that the resulting model worked well in determining the sentiment of reviews with an accuracy score of 73.5%.

## مستخلص البحث

قطرو الندى، نور جيهان. ٢٠٢٣. تحليل المشاعر لمراجعات مستخدمي تطبيق Zoom باستخدام طريقة الساذجة بايز (Naive Bayes). بحث جامعي. قسم الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف: (١) هيشام فاهمي، الماجستير. (٢) إيرنا هيراواتي، الماجستير.

**الكلمات المفتاحية:** المعالجة المسبقة للبيانات ، Naive Bayes ، التحليل المشاعر ، Zoom.

تم تنفيذ القيود الاجتماعية الواسعة لمنع انتشار مرض فيروس كورونا. لذلك ، يجب أن تعقد الأنشطة في القطاعات التعليمية والاقتصادية والاجتماعية والثقافية عبر الإنترنت. أحد تطبيقات مؤتمرات الفيديو المستخدمة بشكل متكرر هو Zoom. إلى جانب الزيادة في مستخدمي تطبيق التكبير، يتم تقديم الاقتراحات والانتقادات من قبل المستخدمين من خلال المراجعات للحصول على التجربة الجيدة. دراسة حسابية هي دراسة التي تدرس الآراء والسلوك والمشاعر تجاه كائن ما. يمكن بعد ذلك استخدام نتائج تحليل المشاعر كاعتبار في اتخاذ القرارات. تهدف هذه الدراسة إلى تحديد تحليل المشاعر لدى مستخدمي تطبيق Zoom باستخدام Naive Bayes. تم الحصول على المراجعات المستخدمة في هذه الدراسة من موقع متجر جوجل بلاي بعدد ٢٠٠ بيانات تتكون من ٥٠ مراجعة بتصنيفات ١، ٢، ٤ و ٥. أولاً ، يجب أن تمر مجموعة بيانات المراجعات بخطوة المعالجة المسبقة عن طريق إزالة علامات الترقيم والأرقام والكلمات التي لا معنى لها والتي لا تؤثر على دقة تحليل المشاعر. بعد ذلك، سيتم اختبار مجموعة البيانات عن طريق تصنيف المشاعر باستخدام Naive Bayes. أظهرت النتائج أن النموذج الناتج يعمل بشكل جيد في تحديد اتجاه المراجعات بدرجة دقة ٧٣,٥%.

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Penyebaran *Corona Virus Disease 2019* (COVID-19) ditetapkan sebagai *Global Pandemic* oleh organisasi kesehatan dunia *World Health Organization* (WHO) pada 11 Maret 2020. Pandemi COVID-19 juga ditetapkan sebagai bencana nonalam nasional berdasarkan Keputusan Presiden Nomor 12 Tahun 2020 tentang Penetapan Bencana Nonalam Penyebaran *Corona Virus Disease 2019* (COVID-19) sebagai Bencana Nasional. COVID-19 disebabkan oleh virus SARS-Cov-2. Penularan COVID-19 banyak terjadi pada manusia yang berinteraksi secara langsung dengan jarak dekat melalui *droplet* atau partikel cairan berukuran kecil yang terlepas ketika sedang berbicara, batuk, bersin atau bernapas (Aeni, 2021).

Salah satu upaya pencegahan penyebaran COVID-19 oleh pemerintah adalah dengan menyelenggarakan Pembatasan Sosial Berskala Besar (PSBB) melalui terbitnya Permenkes Nomor 9 tahun 2020 yang berisi panduan PSBB dalam rangka percepatan penanganan pandemi COVID-19 (Wahidah, dkk., 2020). Pelaksanaan PSBB meliputi antara lain: peliburan sekolah dan tempat kerja, pembatasan kegiatan keagamaan, sosial dan budaya, pembatasan moda transportasi dan pembatasan kegiatan di tempat atau fasilitas umum. Salah satu akibat dari pelaksanaan PSBB yaitu terselenggaranya kegiatan pembelajaran, pekerjaan, maupun kegiatan keagamaan, sosial dan budaya lainnya secara daring dengan menggunakan aplikasi *video conference* seperti Zoom, Skype, Microsoft Teams,

dan lain-lain. Aplikasi Zoom merupakan salah satu aplikasi yang paling marak digunakan.

Seiring bertambahnya penggunaan aplikasi Zoom, perbaikan dan perkembangan aplikasi terus dilakukan untuk memenuhi ekspektasi pengguna. Perusahaan dapat menganalisis ulasan yang diberikan oleh pengguna untuk mendapatkan informasi seperti sentimen pengguna terhadap produk, fitur yang disenangi dari produk, maupun fitur yang perlu diperbaiki dari produk. Analisis sentimen merupakan studi komputasi yang mempelajari opini, sikap, serta perasaan terhadap suatu objek (Sumanjaya, dkk., 2022). Dengan menerapkan analisis sentimen, perusahaan dapat menganalisis memperoleh informasi yang diperlukan untuk melakukan perbaikan maupun mengembangkan produk menjadi lebih baik.

Hal tersebut tersirat dalam firman Allah pada surat Al-Hasyr ayat 18 yang berbunyi:

يَا أَيُّهَا الَّذِينَ ءَامَنُوا اتَّقُوا اللَّهَ وَانْتظِرْ نَفْسَ مَا قَدَّمْتُمْ لِغَدٍ ۖ وَاتَّقُوا اللَّهَ ۚ إِنَّ اللَّهَ خَبِيرٌ بِمَا تَعْمَلُونَ

*“Wahai orang-orang yang beriman! Bertakwalah kepada Allah dan hendaklah setiap orang memperhatikan apa yang telah diperbuatnya untuk hari esok (akhirat), dan bertakwalah kepada Allah. Sungguh, Allah Maha teliti terhadap apa yang kamu kerjakan.” (Al-Hasyr: 18)*

Pada ayat tersebut diserukan kepada orang yang beriman untuk bertaqwa kepada Allah dengan melaksanakan segala perintah-Nya dan menjauhi larangan-Nya. Orang yang bertakwa kepada Allah hendaknya senantiasa memperhatikan apakah yang dikerjakannya akan membawa manfaat kepada dirinya di akhirat kelak atau tidak. Selain itu, hendaknya memperhitungkan apakah perbuatannya sesuai dengan perintah Allah atau tidak. Hal tersebut sejalan dengan tujuan dari dilaksanakannya analisis sentimen.

Herlinawati, Yuliani, Faizah, Gata dan Samudi (2020) melakukan analisis sentiment pada ulasan aplikasi Zoom Cloud dari Google Play Store dengan menggunakan *Naïve Bayes* dan *Support Vector Machine (SVM)*. Metode *Naïve Bayes* menghasilkan analisis dengan tingkat akurasi sebesar 74,37% dan sebesar 81,22% untuk metode *Support Vector Machine*. Limbong, Sembiring dan Hartono (2022) melakukan analisis sentimen terhadap ulasan pada *e-commerce* Shopee dengan menggunakan metode *Naïve Bayes* dan *K-Nearest Neighbor (KNN)* yang menghasilkan tingkat akurasi sebesar 91,4% dan 92,8% masing-masing. Analisis sentimen berbasis aspek dilakukan oleh Parasati, Bachtiar dan Setiawan pada tahun 2020 pada ulasan pelanggan Restoran Bakso *President* Malang menggunakan metode *Naïve Bayes Classifier* dengan aspek yang diteliti yaitu makanan, layanan dan atmosfer. Penelitian tersebut menghasilkan tingkat akurasi pada aspek makanan sebesar 88%, aspek layanan sebesar 76% dan aspek atmosfer sebesar 84%. Analisis sentimen dengan menggunakan metode *Naïve Bayes* memberikan hasil dengan keakuratan yang relatif tinggi. Berdasarkan uraian di atas, penulis mengangkat penelitian dengan judul “Analisis Sentimen Pengguna Aplikasi Zoom Menggunakan Metode *Naïve Bayes*”.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah yang diangkat pada penelitian ini adalah bagaimana hasil analisis sentimen dari ulasan pengguna aplikasi Zoom menggunakan metode *Naïve Bayes*.

### 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, tujuan dari penelitian ini yaitu untuk mengetahui hasil analisis sentimen dari ulasan pengguna aplikasi Zoom menggunakan algoritma *Naïve Bayes*.

### 1.4 Manfaat Penelitian

Adapun manfaat dalam penelitian ini adalah sebagai tambahan bahan referensi untuk pengembangan penelitian lebih lanjut pada bidang matematika komputasi khususnya yang berhubungan dengan metode *Naïve Bayes* maupun analisis sentimen.

### 1.5 Batasan Masalah

Agar penelitian lebih terarah, maka permasalahan yang akan dibahas perlu diberi batasan. Penelitian ini menggunakan data ulasan aplikasi Zoom berbahasa Indonesia yang diambil dari situs Google Play Store. Data ulasan yang diambil hanya ulasan yang memiliki *rating* bintang 1, 2, 4 dan 5.

### 1.6 Definisi Istilah

Berdasarkan fokus dan rumusan masalah penelitian, maka uraian definisi dalam penelitian ini sebagai berikut:

Opini : Tanggapan atau pendapat terhadap suatu rangsangan yang dibentuk berdasarkan interpretasi pribadi yang dapat berupa tulisan maupun ungkapan.

Polaritas : Kecenderungan nuansa kalimat apakah positif atau negatif.

*Training Dataset* : Dataset yang digunakan untuk membangun model sesuai dengan algoritma yang digunakan.

*Testing Dataset* : Dataset yang digunakan untuk menguji dan menilai kinerja dari model yang telah dibangun oleh *training* dataset.

*Cross Validation* : Metode untuk membagi dataset menjadi training dan testing dataset.

## BAB II

### KAJIAN TEORI

#### 2.1 Analisis Sentimen

Analisis sentimen adalah studi komputasi yang mempelajari opini, sikap, serta perasaan terhadap suatu objek (Sumanjaya, dkk., 2022). Objek yang dimaksud dapat berupa orang, acara, produk, maupun topik. Analisis sentimen atau yang bisa juga disebut sebagai *opinion mining* merupakan proses memahami, mengekstrak dan mengolah opini untuk memperoleh informasi sentimen yang terkandung di dalamnya (Nurzahputra & Muslim, 2016). Opini adalah pendapat, ide atau pikiran untuk menjelaskan kecenderungan tertentu terhadap suatu perpektif dan bersifat tidak objektif (Ratnawati, 2018). Analisis Sentimen biasanya digunakan untuk mengidentifikasi kecenderungan pasar atau respon terhadap suatu objek produk Tujuan dari analisis sentimen adalah untuk memperoleh informasi berupa polaritas sebuah opini apakah opini tersebut cenderung positif, negatif, maupun netral. Hasil dari analisis tersebut kemudian dijadikan bahan pertimbangan untuk mengambil keputusan.

Analisis sentimen memiliki dapat dilaksanakan dalam beberapa level, yaitu (Syah, dkk., 2017):

1. *Document level sentiment classification*

*Document level sentiment classification* adalah proses klasifikasi sentimen yang mengolah keseluruhan opini untuk menentukan apakah opini tersebut merupakan opini negatif, positif, maupun netral.

## 2. *Sentence level sentiment classification*

*Sentence level sentiment classification* adalah proses klasifikasi sentimen yang bertujuan untuk menentukan apakah sebuah kalimat dalam opini merupakan kalimat beropini negatif, positif, maupun netral. Pada *Sentence level sentiment classification*, terdapat dua tahapan yaitu menentukan apakah sebuah kalimat dalam opini merupakan kalimat objektif atau subjektif. Kemudian kalimat subjektif diklasifikasi menjadi kalimat beropini negatif, positif, atau netral.

## 3. *Feature-based opinion mining and summarization*

*Feature-based opinion mining and summarization* adalah proses klasifikasi sentimen yang bertujuan untuk memperoleh objek/*feature* yang sedang dibicarakan dalam opini.

## 2.2 *Data Preprocessing*

*Data Preprocessing* merupakan langkah awal yang perlu dilakukan sebelum mengolah data. *Data Preprocessing* bertujuan untuk mengolah data dari yang tidak terstruktur menjadi terstruktur untuk memudahkan proses analisis. Tahapan *Data Preprocessing* meliputi:

1. *Data Cleaning*. Pada tahapan ini, karakter selain huruf akan dihapus termasuk tanda baca dan angka.
2. *Case Folding*. Pada tahapan ini, semua huruf dalam dokumen diubah menjadi bentuk *lower text* agar teks dalam bentuk standar.
3. *Tokenization*. Pada tahapan ini, dokumen akan dibagi dengan memotong kata yang dipisahkan oleh spasi sehingga menjadi bentuk tunggal.

4. *Stopword Removal*. Pada tahapan ini, akan dilakukan penghapusan kata-kata yang tidak bermakna seperti kata sambung.
5. *Stemming*. Pada tahapan ini, kata-kata pada dokumen akan diubah ke dalam bentuk dasarnya dengan menghilangkan kata imbuhan (Yutika, dkk., 2021).

### 2.3 *Naïve Bayes*

*Naive Bayes* adalah salah satu metode klasifikasi paling populer dalam studi *machine learning*. Teknik ini merupakan teknik yang sederhana untuk dilakukan namun memberikan hasil dengan tingkat akurasi yang tinggi terutama dalam pengolahan teks (Ling, dkk., 2014). *Naive Bayes* adalah metode yang didasarkan pada dua asumsi. Pertama, atributnya dianggap tidak tergantung dengan atribut lain. Pada penerapannya, *Naive Bayes* mengasumsikan bahwa keberadaan variabel lain tidak berhubungan dengan keberadaan variabel lainnya (Saputro & Sari, 2019). Kedua, setiap atribut memiliki keunggulan yang sama. Berdasarkan asumsi tersebut, *Naive Bayes* bekerja dengan mempelajari probabilitas yang ada di masa lalu untuk menentukan probabilitas di masa mendatang (Hidayatillah, dkk., 2019).

Untuk memahami ini diperlukan pemahaman tentang teorema Bayes. Teorema Bayes adalah algoritma yang bekerja berdasarkan kondisi probabilitas yang mengukur probabilitas terjadinya sesuatu berdasarkan peristiwa yang telah terjadi di masa lalu. Persamaan probabilitas yang mendasari *Naive Bayes* adalah:

$$P(c|x) = \frac{P(x|c) \times P(c)}{P(x)}$$

dimana  $P(c|x)$  adalah nilai probabilitas yang kita cari,  $P(c)$  adalah kelas probabilitas yang berdasarkan hipotesis,  $P(x|c)$  adalah prediktor probabilitas yang berdasarkan pada kelas yang diberikan dan  $P(x)$  adalah prediktor probabilitas. *Naïve Bayes*

adalah salah satu metode yang cocok dalam *data mining* yang mengkombinasikan teorema Bayes dengan *Maximum A Posteriori* (MAP). Rumus MAP dapat ditulis sebagai berikut:

$$CMAP = \operatorname{argmax} P(X|c) \times P(c)$$

dimana  $c$  adalah kelas target, dan  $P(X|c)$  probabilitas dari kelas yang berdasarkan hipotesis, sementara  $P(C)$  adalah probabilitas prior. Tugasnya adalah untuk mencari kelas yang benar dari setiap sampel (Hidayatillah, dkk., 2019)

#### 2.4 *Confusion Matrix*

*Confusion Matrix* adalah salah satu metode yang digunakan untuk menghitung nilai akurasi dari hasil analisis sentimen. *Confusion Matrix* terdiri dari data hasil prediksi dan data aktual. *Confusion Matrix* digambarkan dengan tabel yang menyatakan jumlah data hasil uji yang benar diklasifikasikan dan data hasil uji yang salah diklasifikasikan (Rahman, dkk., 2017).

**Tabel 2.1** *Confusion Matrix*

<i>Confusion Matrix</i>		<i>Actual Data</i>	
		<i>Positive</i>	<i>Negative</i>
<i>Prediction Result</i>	<i>Positive</i>	<i>True Positive</i>	<i>False Negative</i>
	<i>Negative</i>	<i>False Positive</i>	<i>True Negative</i>

Berdasarkan tabel *Confusion Matrix* diatas, diketahui:

- a. ***True Positive (TP)*** adalah jumlah data aktual positif yang berhasil diklasifikasikan oleh model sebagai hasil prediksi positif.
- b. ***False Positive (FP)*** adalah jumlah data aktual negatif yang terklasifikasikan oleh model sebagai hasil prediksi positif.

- c. *False Negative (FN)* adalah jumlah data aktual positif yang terklasifikasikan oleh model sebagai hasil prediksi negatif.
- d. *True Negative (TN)* adalah jumlah data aktual negatif yang berhasil diklasifikasikan oleh model sebagai hasil prediksi negative.

Melalui 4 data tersebut, dapat diperoleh nilai lain yang bisa digunakan untuk menguji kinerja metode yang telah dibentuk, yaitu (Normawati & Prayogi, 2021) :

1. *Accuracy*. Akurasi menggambarkan seberapa akurat model yang telah dibentuk. Nilai akurasi diperoleh dari persamaan berikut:

$$Accuracy = \frac{TP + TN}{Jumlah\ data}$$

2. *Precision*. *Precision* menggambarkan kemampuan model dalam memanggil kelas yang diminta dibandingkan dengan keseluruhan hasil prediksi dari kelas tersebut. Misal, kemampuan dalam memprediksi data aktual positif dari semua hasil prediksi kelas positif. Nilai *precision* dapat dihitung melalui persamaan berikut:

$$Precision = \frac{TP}{TP + FP}$$

3. *Recall*. *Recall* menggambarkan kemampuan model dalam memanggil kelas yang diminta dari keseluruhan data dari kelas tersebut. Misal, kemampuan model untuk memprediksi data aktual positif dari keseluruhan jumlah data aktual positif pada dataset. Nilai *recall* dapat dihitung melalui persamaan berikut:

$$Recall = \frac{TP}{TP + FN}$$

4. *F1-score*. *F1-score* adalah rata-rata harmonik dari *precision* dan *Recall*.

Nilai *F1-score* dapat dihitung melalui persamaan berikut:

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

## 2.5 Kajian Keislaman

Muhasabah adalah merenungkan dan memperhatikan hal-hal baik maupun buruk yang telah dilakukan, termasuk memperhatikan niat dan tujuan dari suatu perbuatan, serta menghitung untung dan rugi dari tindakan tersebut. Muhasabah diri identik dengan menilai diri sendiri, mengevaluasi, atau introspeksi diri dengan mengacu kepada al-Quran dan hadis Nabi sebagai dasar penilaian, bukan berdasarkan keinginan diri sendiri. Muhasabah adalah salah satu cara untuk memperbaiki, melatih, menyucikan, dan membersihkan hati. Setiap umat muslim dituntut untuk senantiasa melakukan muhasabah berdasarkan firman Allah dalam surah al-Hasyr ayat 18:

يَا أَيُّهَا الَّذِينَ آمَنُوا اتَّقُوا اللَّهَ وَلْتَنْظُرْ نَفْسٌ مَّا قَدَّمَتْ لِغَدٍ وَاتَّقُوا اللَّهَ إِنَّ اللَّهَ خَبِيرٌ بِمَا تَعْمَلُونَ

*“Wahai orang-orang yang beriman! Bertakwalah kepada Allah dan hendaklah setiap orang memperhatikan apa yang telah diperbuatnya untuk hari esok (akhirat), dan bertakwalah kepada Allah. Sungguh, Allah Maha Teliti terhadap apa yang kamu kerjakan.” (Al-Hasyr: 18)*

Berdasarkan Tafsir Ibnu Katsir, ayat tersebut menjelaskan bahwa Allah menyerukan agar umat manusia menghisab diri sendiri sebelum dihisab oleh Allah pada Hari Pembalasan. Ayat tersebut merupakan isyarat agar setiap manusia membuat perhitungan terhadap amal yang dilakukannya yang telah disimpan sebagai bekal untuk hari kiamat esok. Kemudian ditegaskan kembali kepada umat

umat manusia untuk bertaqwa kepada Allah karena sesungguhnya Allah Maha Mengetahui segala perbuatan makhluk-Nya tanpa terkecuali.

Analisis sentimen merupakan proses memahami, mengekstrak dan mengolah opini untuk memperoleh informasi sentimen yang terkandung di dalamnya. Analisis sentiment bertujuan untuk memperoleh informasi berupa polaritas sebuah opini apakah opini tersebut cenderung positif, negatif, maupun netral. Hasil dari analisis tersebut kemudian dijadikan bahan pertimbangan untuk mengambil keputusan, begitu pula dengan proses muhasabah diri.

Muhasabah diri meliputi evaluasi terhadap segala hal yang dilakukan oleh seluruh anggota tubuh, termasuk pemanfaatan usia yang telah diberikan oleh Allah. Melakukan introspeksi diri secara rutin perlu dilakukan agar senantiasa terhindar dari keburukan diri dan mampu mengendalikan diri masing-masing. Hal tersebut meliputi rasa syukur kepada Allah Swt. dan juga rasa malu atas segala dosa dan kesalahan yang telah diperbuat.

Rasulullah bersabda dalam sebuah hadits,

عن أنس بن مالك رضي الله عنه قال: قال رسول الله صلى الله عليه وسلم: كلُّ بني آدم خَطَّاءٌ، وخَيْرُ الخَطَّائِينَ التَّوَابُونَ.

*“Dari Anas ia berkata, Rasulullah bersabda, “Setiap anak Adam mempunyai kesalahan dan sebaik-baiknya orang yang berbuat kesalahan adalah yang bertaubat.” (HR. At-Tirmidzi dan Ibnu Majah)*

Hadits yang diriwayatkan oleh Imam At-Tirmidzi ini menjelaskan bahwa setiap manusia tidaklah luput dari kesalahan. Meskipun demikian, hal tersebut bukanlah sebuah alasan untuk tetap melakukan kesalahan dan bertingkah sewenang-wenang tanpa mengintrospeksi diri. Hal tersebut juga bukan sebuah alasan untuk bersikap pesimis terhadap Allah, karena Allah membuka pintu taubat

bagi hamba-Nya yang menyadari dan menyesali kesalahannya. Sehingga sangat dianjurkan bagi umat muslim untuk senantiasa mengintrospeksi diri dan bertaubat setelah melakukan kesalahan. Seseorang yang cerdas akalnya, akan enggan untuk mengulangi kesalahan yang sama.

Muhasabah adalah jalan orang-orang yang beriman. Seorang Mukmin yang bertaqwa kepada Allah akan selalu bermuhasabah. Orang yang senantiasa bermuhasabah tidak akan dengan mudahnya tergelincir dari jalan kebaikan karena mereka selalu menyibukkan diri dengan menaati perintah Allah dan mencela diri atas kelalaiannya terhadap perintah Allah. Maka evaluasi diri menjadi suatu yang penting untuk tetap menjaga keseimbangan diri agar selalu berada di jalan yang benar.

## BAB III

### METODE PENELITIAN

#### 3.1 Jenis Penelitian

Pada penelitian ini, pendekatan yang digunakan adalah pendekatan studi kasus kuantitatif. Studi kasus deskriptif dilakukan dengan mengumpulkan data ulasan pengguna aplikasi Zoom pada Google Play Store Review. Data yang diperoleh kemudian dianalisis dan disusun sesuai dengan kebutuhan.

#### 3.2 Data dan Sumber Data

Pada Penelitian ini, data yang digunakan adalah data berjenis teks yang berupa ulasan pengguna aplikasi Zoom. Data didapatkan dari situs Google Play Store yang didapatkan melalui *web scraping*. Data yang digunakan adalah ulasan dari pengguna dengan *rating* bintang 1, *rating* bintang 2, *rating* bintang 4 dan *rating* bintang 5 masing-masing berjumlah 50 ulasan. Ulasan dengan *rating* bintang 1 dan *rating* bintang 2 akan diberi label negative. Sedangkan ulasan dengan *rating* bintang 4 dan *rating* bintang 5 akan diberi label positif.

**Tabel 3.2** Sampel Data Ulasan Pengguna Zoom

<b>Rating</b>	<b>Ulasan</b>
1	sangat tidak bagus mau di daftar selalu ngak bisa
1	Suara sering hilang dan suara sering berdeper kayak robot terus suka keluar sendiri dari rapat
1	Sekarang suaranya hilang tinggal gambarnya yang muncul
1	Gak bisa ganti background di zoom android
2	Bug parah ketika end zoom malah stuck di black atau blank screen. Padahal sekelas flagship samsung. Di versi sebelumnya tidak seperti ini

Lanjutan **Tabel 3.2** Sampel Data Ulasan Pengguna *Zoom*

<b>Rating</b>	<b>Ulasan</b>
2	Setelah diupdate, chat di zoom tidak muncul dan sering hilang.
2	Zoom tidk ada virtual backgroundnya di HP.. Beda sama zoom dulu
4	Membantu saat tak bisa rapat tatap muka
4	Salah satu alternatif untuk berkomunikasi dan melakukan pertemuan secara publik
4	Jaringan bagus, suara tdk terputus putus terimakasih
4	Sejak menggunakan zoom meeting komunikasi dengan rekan ketja semakin inten, mudah berbagi informasi
5	Memudahkan kita untuk melakukan kegiatan yang membutuhkan aplikasi rapat bila jarak tempuh yang jauh.. tidak repot dgn adanya zoom, kegiatan lebih terasa lebih ringan dan mudah
5	Terima kasih aplikasi zoom telah menemaniku kuliah selama masa pandemi dan sekarang alhamdulillah masa pandemi berakhir, dan aplikasi ini memudahkan kita untuk bersilaturahmi jarak jauh
5	Aplikasinya sangat berguna sekali, bermanfaat, no iklan2, sip, bintang 5
5	Lancar, jelas dan terang gambarnya

### 3.3 Tahapan Penelitian

Tahapan-tahapan yang dilalui untuk menganalisis sentimen ulasan pengguna aplikasi Zoom adalah sebagai berikut:

#### A. *Data Preprocessing.*

Data ulasan pengguna aplikasi *zoom* yang diperoleh berbentuk teks yang terdiri dari tidak hanya karakter huruf saja, melainkan bercampur dengan angka, tanda baca, emotikon, dan lain-lain. Data teks merupakan data yang tidak berstruktur, sehingga perlu dilakukan *preprocessing* terlebih dahulu sebelum kemudian diolah. *Data preprocessing* akan mengubah dataset yang masih

berupa teks dan tidak berstruktur menjadi data yang terstruktur. Proses yang dilalui dalam tahapan *data preprocessing* antara lain:

1. *Data Cleaning*. Karakter selain huruf, seperti angka dan tanda baca akan dihilangkan.
2. *Case Folding*. Setelah menyisakan hanya karakter berupa huruf saja, *dataset* kemudian diubah ke dalam bentuk *lowercase*.
3. *Tokenizing*. Proses selanjutnya yaitu *tokenizing*, dimana masing-masing dokumen akan dipisahkan menjadi bentuk tunggal dengan pemisah spasi.
4. *Stopwords Removal*. Data-data tunggal kemudian akan di-*filter* untuk menghilangkan kata-kata yang tidak memiliki makna, ataupun kata yang terlalu sering disebutkan sehingga menghilangkan nilai sentiment yang dimiliki.
5. *Stemming*. Kata-kata yang melewati proses *stopword removal* kemudian akan diubah kedalam bentuk dasar dengan menghilangkan imbuhan depan maupun imbuhan belakang yang dimiliki.

B. *Data Training*.

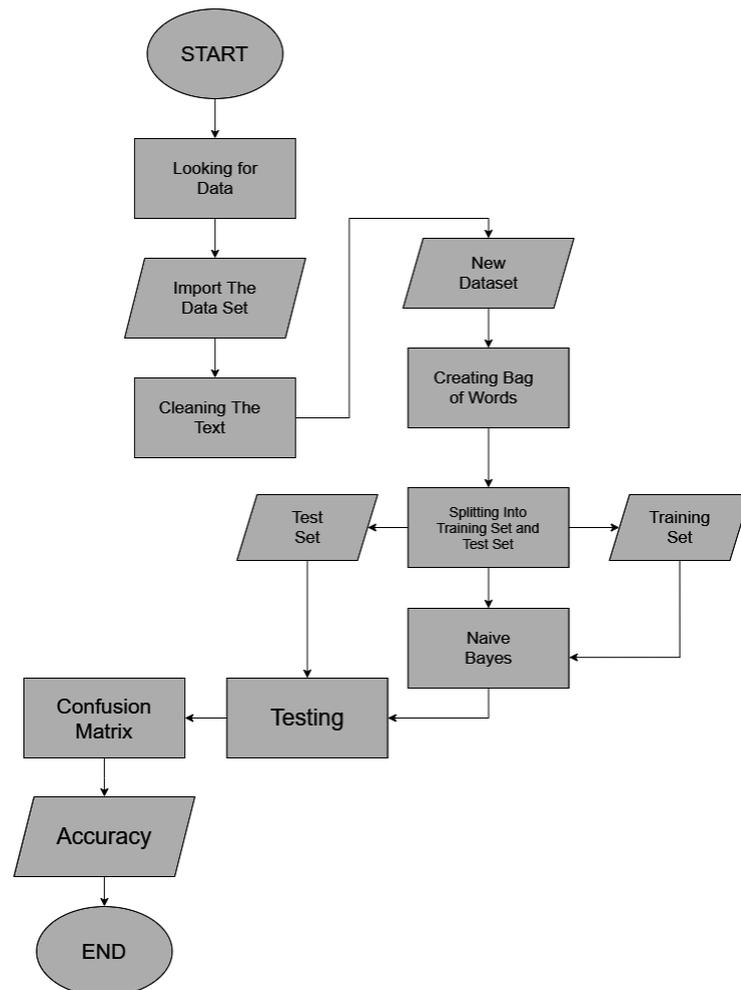
Setelah melewati tahapan *preprocessing*, *dataset* kemudian akan dibagi menjadi dua kelompok yaitu, *training dataset* dan *testing dataset*. *Training dataset* akan digunakan untuk membentuk model *classifier* dengan pendekatan *Bag of Words*. Metode ini akan mengubah data teks ke dalam bentuk vector dengan menghitung frekuensi kemunculan kata dari tiap ulasan.

### C. *Data Testing.*

Pada tahapan *data testing*, model yang terbentuk dari hasil *training dataset* kemudian akan diaplikasikan pada *testing dataset* untuk melihat tingkat akurasi dari model yang telah dibentuk.

### D. *Evaluation.*

Hasil klasifikasi yang diperoleh dari penerapan model pada *testing dataset* kemudian akan dihitung tingkat akurasinya. Penghitungan akurasi dilakukan dengan memanfaatkan *confusion matrix*. Selain itu, proses yang sama akan diterapkan lagi pada kelompok dataset lainnya menggunakan *cross validation* sehingga diperoleh nilai yang dapat mewakili tingkat akurasi secara umum.



**Gambar 3.1** Flowchart Tahapan Analisis Sentimen

## BAB IV

### PEMBAHASAN

#### 4.1 Hasil Tahapan *Data Preprocessing*

##### 4.1.1 *Data Cleaning dan Case Folding*

**Tabel 4.1** Proses *Data Cleaning* dan *Case Folding*

No	Sebelum proses <i>data cleaning</i>	Setelah proses <i>data cleaning</i>
1	Terimakasih sudah bantu aku mengerjakan tugas via zoom. Diskusi makin lancar	terimakasih sudah bantu aku mengerjakan tugas via zoom diskusi makin lancar
2	Tidak ada opsi untuk pakai background. Hpku samsung a30s	tidak ada opsi untuk pakai background hpku samsung a s
3	Setelah diupdate, chat di zoom tidak muncul dan sering hilang.	setelah diupdate chat di zoom tidak muncul dan sering hilang

Tabel 4.1 menunjukkan hasil proses *data cleaning* dan *case folding*. Tanda baca ‘titik’ pada contoh ulasan ke-1 dan tanda baca ‘koma’ pada ulasan ke-3 dihilangkan. Angka ‘30’ pada ulasan ke-2 juga dihilangkan.

##### 4.1.2 *Tokenization*

**Tabel 4.2** Proses *Tokenization*

No	Sebelum proses <i>tokenization</i>	Setelah proses <i>tokenization</i>
1	setelah diupdate chat di zoom tidak muncul dan sering hilang	['setelah', 'diupdate', 'chat', 'di', 'zoom', 'tidak', 'muncul', 'dan', 'sering', 'hilang']
2	tidak ada opsi untuk pakai background hpku samsung a s	['tidak', 'ada', 'opsi', 'untuk', 'pakai', 'background', 'hpku', 'samsung', 'a', 's']
3	Terimakasih sudah bantu aku mengerjakan tugas via zoom. Diskusi makin lancar	['terimakasih', 'sudah', 'bantu', 'aku', 'mengerjakan', 'tugas', 'via', 'zoom', 'diskusi', 'makin', 'lancar']

Tabel 4.2 menunjukkan hasil dari proses *tokenization*. Proses ini memisahkan dokumen ulasan menjadi bentuk tunggal dengan pemisah spasi.

#### 4.1.3 Stopword Removal

Tabel 4.3 Proses *Stopword Removal*

No	Sebelum proses <i>stopword removal</i>	Setelah proses <i>stopword removal</i>
1	['setelah', 'diupdate', 'chat', 'di', 'zoom', 'tidak', 'muncul', 'dan', 'sering', 'hilang']	['diupdate', 'chat', 'zoom', 'tidak', 'muncul', 'hilang']
2	['tidak', 'ada', 'opsi', 'untuk', 'pakai', 'background', 'hpku', 'samsung', 'a', 's']	['tidak', 'opsi', 'pakai', 'background', 'hpku', 'samsung', 'a', 's']
3	['terimakasih', 'sudah', 'bantu', 'aku', 'mengerjakan', 'tugas', 'via', 'zoom', 'diskusi', 'makin', 'lancar']	['terimakasih', 'bantu', 'mengerjakan', 'tugas', 'via', 'zoom', 'diskusi', 'lancar']

Tabel 4.3 menunjukkan hasil dari proses *stopword removal*. *Stopword* yang berhasil ter-*filter* dari contoh dokumen ulasan di atas antara lain: setelah, di, dan, ada, untuk, sudah, aku, dan makin. Kata-kata tersebut tidak memiliki makna yang berpengaruh terhadap tingkat akurasi analisis sentiment.

#### 4.1.4 Stemming

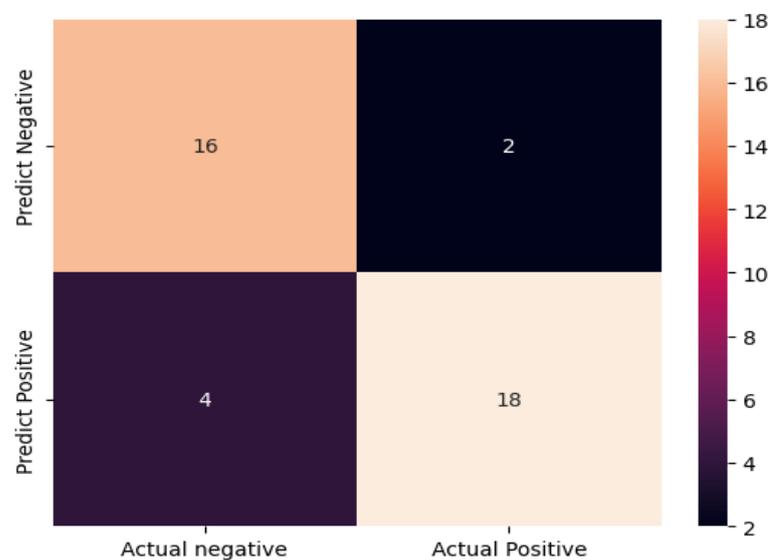
Tabel 4.4 Proses *Stemming*

No	Sebelum proses <i>stemming</i>	Setelah proses <i>stemming</i>
1	['diupdate', 'chat', 'zoom', 'tidak', 'muncul', 'hilang']	['update', 'chat', 'zoom', 'tidak', 'muncul', 'hilang']
2	['tidak', 'opsi', 'pakai', 'background', 'hpku', 'samsung', 'a', 's']	['tidak', 'opsi', 'pakai', 'background', 'hpku', 'samsung', 'a', 's']
3	['terimakasih', 'bantu', 'mengerjakan', 'tugas', 'via', 'zoom', 'diskusi', 'lancar']	['terimakasih', 'aku', 'kerja', 'tugas', 'via', 'zoom', 'diskusi', 'lancar']

Tabel 4.4 menunjukkan hasil dari proses *stemming*. Proses *stemming* memberikan hasil berupa kata dasar dari kata yang terkandung dalam dokumen ulasan. Hasil dari proses *stemming* dapat dilihat pada contoh ulasan ke-1, yaitu kata ‘diupdate’ yang diubah menjadi bentuk kata dasar yaitu ‘update’. Hasil yang lain terdapat pada contoh ulasan ke-3, yaitu kata ‘mengerjakan’ yang diubah ke dalam bentuk dasar menjadi ‘kerja’.

#### 4.2 Hasil Analisis Sentimen Pengguna Aplikasi Zoom

Dalam pengujian pada 40 data menggunakan *Naive Bayes*, diperoleh nilai akurasi sebesar 85%. Evaluasi ini ditunjukkan dari *confusion matrix* dengan *True Positive* sebanyak 16, *False Positive* sebanyak 4, *False Negative* sebanyak 2, dan *True Negative* sebanyak 18, seperti ditunjukkan pada Gambar 4.1.



**Gambar 4.1** *Confusion Matrix* Hasil Analisis Sentimen

Perhitungan nilai akurasi sebagai berikut:

$$accuracy = \frac{TP + TN}{Jumlah\ Data} = \frac{18 + 16}{40} = \frac{34}{40} = 0.85 = 85\%$$

Dengan akurasi sebesar 0.85, dapat disimpulkan bahwa model klasifikasi memberikan prediksi yang akurat dalam mengklasifikasikan sentimen pada data pengujian. True Positive sebanyak 18 menunjukkan bahwa model berhasil mengidentifikasi 16 data yang sebenarnya positif dengan benar. True Negative sebanyak 16 menunjukkan bahwa model juga berhasil mengidentifikasi 16 data yang sebenarnya negatif dengan benar. Namun, terdapat juga beberapa prediksi yang tidak akurat. False Positive sebanyak 4 menunjukkan bahwa model mengklasifikasikan 4 data yang sebenarnya negatif sebagai positif secara keliru. Sementara itu, False Negative sebanyak 2 menunjukkan bahwa model mengklasifikasikan 2 data yang sebenarnya positif sebagai negatif secara keliru. Untuk performa metode klasifikasi dari setiap kelas dapat dilihat dari nilai *precision*, *recall*, dan *f1-score* pada setiap kelas di tabel 4.5.

**Tabel 4.5** Hasil Analisis Sentimen Tiap Kelas

	<i>precision</i>	<i>recall</i>	<i>f1-score</i>
<b>Positif</b>	0,82	0,90	0,86
<b>Negatif</b>	0,89	0,80	0,84

Perhitungan nilai metrik pada kelas Positif sebagai berikut:

$$Precision_{positif} = \frac{TP}{TP + FP} = \frac{18}{18 + 4} = \frac{18}{22} = 0,82$$

$$Recall_{positif} = \frac{TP}{TP + FN} = \frac{18}{18 + 2} = \frac{18}{20} = 0,90$$

$$f1 - score_{positif} = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,82 \times 0,90}{0,82 + 0,90} = 0,86$$

Perhitungan nilai metrik pada kelas Positif sebagai berikut:

$$Precision_{negatif} = \frac{TN}{TN + FN} = \frac{16}{16 + 2} = \frac{16}{18} = 0,89$$

$$Recall_{negatif} = \frac{TN}{TN + FP} = \frac{16}{16 + 4} = \frac{16}{20} = 0,80$$

$$f1 - score_{negatif} = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,89 \times 0,80}{0,89 + 0,80} = 0,84$$

Hasil dari evaluasi model dapat dilihat dari nilai *precision* dan *recall* di setiap kelasnya. Dapat dikatakan tingkat kemampuan model dalam mencari ketepatan suatu informasi yang diminta oleh pengguna adalah sebesar 82% untuk kelas positif dan 80% untuk kelas negatif. Sedangkan kemampuan suatu model dalam menemukan kembali suatu informasi adalah sebesar 90% untuk kelas positif dan 89% untuk kelas negatif. Hasil dari nilai *precision*, *recall*, dan *f1-score* mempunyai nilai sebesar 0-1. Semakin tinggi nilainya maka semakin baik.

Proses yang sama kemudian diterapkan pada dataset dengan menggunakan *10-fold cross validation* dan memberikan hasil akurasi yang disajikan pada tabel 4.6.

**Tabel 4.6** Hasil Analisis Sentimen dengan *10-fold cross validation*

Fold ke-	1	2	3	4	5	6	7	8	9	10
Akurasi	0.8	0.65	0.8	0.75	0.7	0.75	0.8	0.75	0.8	0.55

Rata-rata akurasi dari keseluruhan *cross-validation* adalah 0.735, menunjukkan bahwa model *Naive Bayes* dengan metode *Bag of Words* dalam klasifikasi sentimen pada data ulasan pengguna aplikasi Zoom memberikan kinerja yang cukup baik. Terdapat variasi dalam akurasi antara setiap iterasi *cross-validation*, menandakan adanya perbedaan dalam kemampuan model untuk

mengklasifikasikan sentimen pada subset data yang berbeda. Hasil ini memberikan wawasan kepada pengembang atau pemilik aplikasi Zoom tentang sentimen umum pengguna berdasarkan ulasan yang ada. Meskipun akurasi menjadi metrik utama dalam evaluasi, penting juga untuk mempertimbangkan metrik evaluasi lain seperti *precision*, *recall*, dan *F1-score* untuk mendapatkan pemahaman yang lebih komprehensif tentang kinerja model.

Dengan memahami sentimen pengguna, langkah-langkah dapat diambil untuk meningkatkan pengalaman pengguna dan mengatasi masalah yang mungkin timbul. Namun, dengan melihat variasi dalam akurasi antar iterasi, ada kemungkinan bahwa model dapat ditingkatkan dengan penyesuaian lebih lanjut atau menggunakan teknik pemrosesan teks yang lebih canggih.

### **4.3 Analisis Sentimen dalam Pandangan Islam**

Pada dasarnya, analisis sentimen adalah studi komputasi untuk mengekspresikan opini, sentimen, sikap, emosi, evaluasi, penilaian atau pandangan yang terdapat pada suatu teks. Menurut Kamus Besar Bahasa Indonesia (KBBI) sentimen merupakan perasaan hati atau pandangan yang didasarkan pada perasaan yang berlebihan terhadap sesuatu. Jadi, sentimen bisa dikatakan sebagai proses memahami, mengekstrak dan mengolah opini untuk memperoleh informasi sentimen yang terkandung di dalamnya. Pada zaman sekarang ini begitu banyak permasalahan setiap individu yang dikemukakan oleh masyarakat di sosial media. Untuk itu dibutuhkan analisis sentimen untuk memperoleh informasi berupa polaritas sebuah opini apakah opini tersebut cenderung positif, negatif, maupun

netral. Hasil dari analisis tersebut kemudian dijadikan bahan pertimbangan untuk mengambil keputusan. Konsep yang sama juga dimiliki oleh proses muhasabah.

Kata muhasabah berasal dari kata *hasaba-yuhasibu* yang bentuk dasarnya adalah *hasaba-yahsubu* yang artinya menghitung. Sedangkan dalam kamus bahasa Arab-Indonesia, kata muhasabah berarti perhitungan atau introspeksi (Munawwir, 1997). Muhasabah diri adalah bentuk mengintrospeksi, meneliti, atau mengevaluasi diri dengan memperhitungkan segala perbuatan yang telah dilakukan termasuk niat dan tujuan, serta apakah perbuatan yang dilakukan tersebut memberikan manfaat ataukah *mudhorot* bagi dirinya maupun orang lain (Syafri, Qotadah, & Achmad, 2020). Dengan kata lain, Islam mengajarkan bahwa setiap manusia memiliki hak untuk memperbaiki hati dan itu merupakan hal yang tidak dapat dipisahkan dari potensi sekaligus perintah Allah Swt agar manusia senantiasa beriman dan bertaqwa kepada Allah Swt, yaitu dengan melaksanakan perintah-perintah dan menjauhi larangan-larangan-Nya. Termasuk melaksanakan perintah-perintah Allah adalah memurnikan ketaatan dan menundukkan diri hanya kepada-Nya saja, tidaklah ada sedikit pun unsur syirik di dalamnya, melaksanakan ibadah yang diwajibkan Nya dan mengadakan hubungan baik sesama manusia.

Muhasabah diri meliputi evaluasi terhadap segala hal yang dilakukan oleh seluruh anggota tubuh, termasuk usia yang telah diberikan oleh Allah Swt. Rasulullah bersabda dalam hadits yang diriwayatkan oleh Imam at-Tirmidzi yang berbunyi,

الْكَيْسُ مَنْ دَانَ نَفْسَهُ وَعَمِلَ لِمَا بَعْدَ الْمَوْتِ وَالْعَاجِزُ مَنْ أَتْبَعَ نَفْسَهُ هَوَاهَا وَتَمَنَّى عَلَى اللَّهِ

*Artinya: "Orang yang cerdas adalah orang yang mau bermuhasabah dan beramal untuk hari kiamat. Sedangkan orang yang bodoh adalah orang yang selalu mengikuti hawa nafsunya serta berharap sesuatu terhadap Allah (tanpa disertai usaha)" (HR. At-Tirmidzi)*

Orang yang senantiasa bermuhasabah dianggap sebagai orang cerdas oleh Nabi Muhammad Saw. karena semasa hidupnya, mereka memanfaatkan usia yang diberikan oleh Allah Swt. menyiapkan bekal untuk kehidupan setelah mati. Menyisihkan waktu untuk mengevaluasi dan memperhatikan setiap langkah yang diambil supaya hanya memberikan manfaat dan bukan *mudhorot* sebagai bentuk rasa syukur kepada Allah Swt. Saat melakukan kesalahan, mereka akan menyesali perbuatan yang telah dilakukan tersebut dan segera bertaubat. Selain itu, orang yang sering bermuhasabah mampu belajar dari kesalahan sehingga dapat menghindari mengulangi kesalahan yang sama. Waktu yang diberikan oleh Allah Swt. dimanfaatkan semaksimal mungkin untuk mempelajari dan mengevaluasi kesalahan dan tindakan yang dilakukan untuk terus memperbaiki diri.

Sedang orang yang dianggap bodoh oleh Nabi Muhammad Saw. adalah orang yang terlena oleh hawa nafsunya sendiri. Mereka akan lalai akan perintah Allah dan menyia-nyiakan waktunya tanpa menyiapkan bekal untuk hari esok. Orang yang tidak bermuhasabah bersikap acuh terhadap tindakan yang diambilnya tanpa memedulikan akibat dari perbuatannya. Mereka tidak merenungkan kembali perbuatan-perbuatan yang telah dilakukan yang kemudian akan mengakibatkan hilangnya kesempatan untuk berubah menjadi pribadi yang lebih baik.

Perbandingan orang bermuhasabah dengan orang yang acuh akan akibat dari perbuatannya ditunjukkan dalam sabda Rasulullah,

مَنْ كَانَ يَوْمُهُ حَيْرًا مِنْ أَمْسِهِ فَهُوَ رَاحٍ . وَمَنْ كَانَ يَوْمُهُ مِثْلَ أَمْسِهِ فَهُوَ مَعْبُودٌ . وَمَنْ كَانَ  
يَوْمُهُ شَرًّا مِنْ أَمْسِهِ فَهُوَ مَلْعُونٌ

*Artinya: "Barang siapa hari ini lebih baik dari hari kemarin, dialah tergolong orang yang beruntung, Barang siapa yang hari ini sama dengan hari*

*kemarin dialah tergolong orang yang merugi dan Barang siapa yang hari ini lebih buruk dari hari kemarin dialah tergolong orang yang celaka." (HR. Al Hakim).*

Muhasabah adalah jalan orang-orang yang beriman. Seorang Mukmin yang bertaqwa kepada Allah akan selalu bermuhasabah. Orang yang senantiasa bermuhasabah tidak akan dengan mudahnya tergelincir dari jalan kebaikan karena mereka selalu menyibukkan diri dengan menaati perintah Allah dan mencela diri atas kelalaiannya terhadap perintah Allah. Maka evaluasi diri menjadi suatu yang penting untuk tetap menjaga keseimbangan diri agar selalu berada di jalan yang benar.

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan rumusan masalah dan pembahasan pada bab sebelumnya maka diperoleh kesimpulan bahwa model yang dibentuk dari ulasan pengguna aplikasi Zoom bekerja dengan baik. Hal ini ditunjukkan nilai *precision* dan *recall* yang dihasilkan oleh model terhadap *testing data* yaitu 88,9% dan 80% dengan nilai *f1-score* sebesar 84,2%. Selain itu, nilai rata-rata akurasi yang dihasilkan dari iterasi *cross-validation* sebanyak 10 kali adalah sebesar 73,5%.

#### **5.2 Saran**

Berdasarkan hasil pembahasan pada penelitian ini, beberapa saran dari penulis untuk mengembangkan penelitian ini lebih lanjut di antaranya:

1. Melakukan tahapan *data preprocessing* lebih teliti dan menyeluruh, seperti penanganan kata yang disingkat, penggunaan bahasa yang tidak baku, mengoptimalkan penggunaan emoji, dan sebagainya;
2. Meneliti analisis sentiment pada level aspek yang memberikan informasi lebih mendalam mengenai objek yang diteliti dibandingkan dengan analisis sentiment pada level lainnya.

## DAFTAR PUSTAKA

- Aeni, N. (2021). Pandemi COVID-19: Dampak Kesehatan, Ekonomi, dan Sosial. *Jurnal Litbang: Media Informasi Penelitian, Pengembangan dan IPTEK*, 17(1), 17-34.
- Herlinawati, N., Yuliani, Y., Faizah, S., Gata, W., & Samudi, S. (2020). Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine. *CESS (Journal of Computer Engineering, System and Science)*, 5(2), 293-298.
- Hidayatillah, R., Mirwan, M., Hakam, M., & Nugroho, A. (2019). Levels of Political Participation Based on naive Bayes Classifier. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 13(1), 73.
- Limbong, J. J., Sembiring, I., & Hartomo, K. D. (2022). ANALISIS KLASIFIKASI SENTIMEN ULASAN PADA E-COMMERCE SHOPEE BERBASIS WORD CLOUD DENGAN METODE NAIVE BAYES DAN K-NEAREST NEIGHBOR. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 9(2), 347-356.
- Ling, J., Kencana, I. E., & Oka, T. B. (2014). ANALISIS SENTIMEN MENGGUNAKAN METODE NAÏVE BAYES DENGAN SELEKSI FITUR CHI SQUARE. *E-Jurnal Matematika*, 3(3), 92-99.
- Munawwir, A. W. (1997). *Kamus Al-Munawwir Bahasa Arab-Indonesia Terlengkap*. Surabaya: Pustaka Progresif.
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697-711.
- Nurzahputra, A., & Muslim, M. A. (2016, Oktober 10). Analisis Sentimen pada Opini Mahasiswa Menggunakan Natural Language Processing. *Seminar Nasional Ilmu Komputer (SNIK 2016)*.
- Parasati, W., Bachtiar, F. A., & Setiawan, N. Y. (2020). Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Restoran Bakso President Malang dengan Metode Naïve Bayes Classifier. 4(4), 1090-1099.
- Rahman, M., Alamsah, D., Darmawidjadja, M., & Nurma, I. (2017). Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN). 11(1), 36-45.
- Ratnawati, F. (2018). Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter. *JURNAL INOVTEK POLBENG - SERI INFORMATIKA*, 3(1), 50-59.

- Saputro, I. W., & Sari, B. W. (2019). Uji Performa Algoritma Naive Bayes untuk Prediksi Masa Studi Mahasiswa. *Citec Journal*, 6(1), 1-11.
- Sumanjaya, A. A., Indriati, & Ridok, A. (2022). Analisis Sentimen Data Tweets terhadap Penanganan Covid-19 di Indonesia menggunakan Metode Naïve Bayes dan Pemilihan Kata Bersentimen menggunakan Lexicon Based. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6 (4), 1865-1872.
- Syafri, I., Qotadah, H. A., & Achmad, A. D. (2020). Muhasabah Diri Sebagai Media Penanggulangan Perilaku Juvenile Delinquency. *Khazanah Pendidikan Islam*, 2(3), 126-138.
- Syah, A. P., Adiwijaya, & Al faraby, S. (2017). ANALISIS SENTIMEN PADA DATA ULASAN PRODUK TOKO ONLINE DENGAN METODE MAXIMUM ENTROPY. *e-Proceeding of Engineering*, 4(3), 4632-4640.
- Wahidah, I., Septiadi, M. A., Rafqie, M. A., Hartono, N. S., & Athallah, R. (2020). Pandemi Covid-19: Analisis Perencanaan Pemerintah dan Masyarakat dalam Berbagai Upaya Pencegahan. *Jurnal Manajemen dan Organisasi (JMO)*, 11(3), 179-188.
- Yutika, C. H., Adiwijaya, & Al Faraby, S. (2021). Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes. 5(2), 422-430.

## LAMPIRAN

### Downloading Library

```
[ ] !pip install Sastrawi
    !pip install nltk
```

Show hidden output

IMPORT LIBRARIES

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
#matplotlib inline
sns.set_style("whitegrid")
pd.pandas.set_option('display.max_columns', None)
import re
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
import nltk
from nltk.corpus import stopwords
nltk.download('stopwords')
```

### Loading Dataset

```
[ ] dataset = pd.read_csv('review_zoom.csv')
```

```
df = dataset.replace([1,2],0) #review dengan rating *1 dan *2 diberikan label 0 (Negatif)
df = df.replace([4,5],1) #review dengan rating *4 dan *5 diberikan label 1 (Positif)

df = df.rename(columns={"score": "label", "content": "ulasan"}) #rename column
df
```

### Cleaning Data

```
[ ] def clean_text(text):
    return re.sub('[^a-zA-Z]', ' ', text).lower()

df['cleaned'] = df['ulasan'].apply(lambda x:clean_text(x))
df.to_csv('bef_af_cleaning.csv', index=False)
df
```

### Tokenizing Data

```
[ ] def tokenize_text(text):
    tokenized_text = text.split()
    return tokenized_text

df['tokens'] = df['cleaned'].apply(lambda x:tokenize_text(x))

df.to_csv('bef_af_token.csv', index=False)
df
```

## Removing Stopword

```
stopword = nltk.corpus.stopwords.words('indonesian')
stopword.remove('tidak')

def remove_stopwords(text):
    text = [word for word in text if word not in stopword]
    return text

df['stopremov'] = df['tokens'].apply(lambda x: remove_stopwords(x))
df.to_csv('bef_af_stremove.csv', index=False)
df
```

## Stemming

```
def stemming(dokumen):
    factory = StemmerFactory()
    stemmer = factory.create_stemmer()
    dokumen = [stemmer.stem(word) for word in dokumen]
    dokumen = [t for t in dokumen if len(t) > 1]
    dokumen = " ".join(dokumen)
    return(dokumen)

df['stemmed'] = df['tokens'].apply(lambda x: stemming(x))
df.to_csv('bef_af_stemmed.csv', index=False)
df
```

## Modelling

```
corpus = df['stemmed']

from sklearn.feature_extraction.text import CountVectorizer
from sklearn.model_selection import train_test_split
cv = CountVectorizer(max_features = 500)
bow_article = cv.fit(corpus)
count_tokens = bow_article.get_feature_names_out()
article_vect = bow_article.transform(corpus)
df_count_vect = pd.DataFrame(data=article_vect.toarray(), columns=count_tokens)
X = cv.fit_transform(corpus).toarray()
y = df['label']
```

## Training Model

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.20, random_state=0)

#training

from sklearn.naive_bayes import MultinomialNB
classifier = MultinomialNB()
classifier.fit(X_train, y_train)
```

## Testing Model

```
[ ] #testing

y_pred = classifier.predict(X_test)
print(y_pred)
```

## Confusion Matrix

```

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
cm = confusion_matrix(y_test, y_pred)

cm_matrix = pd.DataFrame(data=cm, columns = ['Actual negative', 'Actual Positive'],
                        index = ['Predict Negative', 'Predict Positive'])
sns.heatmap(cm_matrix, annot = True, fmt='d')

TP = cm[0,0]
TN = cm[1,1]
FP = cm[0,1]
FN = cm[1,0]

acc = accuracy_score(y_test, y_pred)
print('Accuracy: ', acc)

precision = TP/float(TP+FP)
print('Precision : {0:0.4f}'.format(precision))

recall = TP / float(TP+FN)
print('Recall or Sensitivity : {0:0.4f}'.format(recall))

fscore = (2*recall*precision/(recall+precision))
print('f1-score: ', fscore)

```

## Classification report

```
[ ] print(classification_report(y_test, y_pred))
```

## Cross Validation

```
[ ] from sklearn.model_selection import cross_val_score

scores = cross_val_score(classifier, X, y, cv= 10)
print('scores: ', scores)
print(scores.mean())
```

## RIWAYAT HIDUP



Nur Jihan Qothrunnada lahir di Kabupaten Mojokerto pada 17 Agustus 2000. Memiliki nama panggilan Jihan, merupakan anak kedua dari tiga bersaudara dari pasangan Bapak Sahid dan Ibu Munifah. Penulis lahir dan tumbuh di Desa Wonosari, Kecamatan Ngoro, Kabupaten Mojokerto.

Pendidikan yang telah ditempuh, antara lain TK Muslimat Wonosari (lulus tahun 2005), MI Naba'ul Ulum Wonosari (lulus tahun 2011), SMPN 1 Ngoro (lulus tahun 2014), dan MAN Mojosari (lulus tahun 2016). Pada tahun yang sama, penulis melanjutkan pendidikan di Universitas Islam Negeri Maulana Malik Ibrahim Malang pada Program Studi Matematika, Fakultas Sains dan Teknologi.

Selama menempuh pendidikan tinggi, penulis pernah menjadi Muharrikah bidang Keamanan Mabna Ummu Salamah (2016), Anggota Himpunan Mahasiswa Jurusan Matematika “Integral” bidang Kematematikaan (2017), dan Teamwork Mathematics English Club (2017-2019). Prestasi yang berhasil dicapai selama menempuh pendidikan tinggi yaitu Best TOEFL Score Achiever dan Best Graduate Student of One Year English Program yang diselenggarakan oleh Pusat Bahasa Universitas Islam Negeri Maulana Malik Ibrahim Malang pada tahun 2018.



KEMENTERIAN AGAMA RI  
UNIVERSITAS ISLAM NEGERI  
MAULANA MALIK IBRAHIM MALANG  
FAKULTAS SAINS DAN TEKNOLOGI  
Jl. Gajayana No. 50 Dinoyo Malang Telp./Fax. (0341)558933

### BUKTI KONSULTASI SKRIPSI

Nama : Nur Jihan Qothrunnada  
NIM : 16610007  
Fakultas / Program Studi : Sains dan Teknologi / Matematika  
Judul Skripsi : Analisis Sentimen Ulasan Pengguna Aplikasi Zoom Menggunakan Metode *Naive Bayes*  
Pembimbing I : Hisyam Fahmi, M.Kom  
Pembimbing II : Erna Herawati, M.Pd

No.	Tanggal	Hal	Tanda Tangan
1.	27 September 2022	Konsultasi Bab I	1.
2.	30 Januari 2023	Konsultasi Bab I, II dan III	2.
3.	23 Februari 2023	Konsultasi Kajian Keislaman	3.
4.	23 Februari 2023	Konsultasi Bab III	4.
5.	20 Juni 2023	Konsultasi Bab IV dan V	5.
6.	20 Juni 2023	Konsultasi Kajian Keislaman	6.
7.	23 Juni 2023	Konsultasi Bab I, II, III, IV dan V	7.
8.	23 Juni 2023	Konsultasi Kajian Keislaman	8.
9.	30 Juni 2023	Konsultasi dan ACC Keseluruhan	9.

Malang, 30 Juni 2023  
Mengetahui,  
Ketua Program Studi Matematika



Dr. Elly Susanti, M. Sc  
NIP. 19741129 200012 2 005