ANALISIS PERBAIKAN KUALITAS LAYANAN PADA APLIKASI *ZOOM CLOUD MEETINGS* BERDASARKAN ULASAN *END-USER* MENGGUNAKAN METODE *TEXT MINING* DAN *FISHBONE DIAGRAM*

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Strata-1
Pada Jurusan Teknik Industri Fakultas Teknologi Industri

Nama : M. Agung Pratama
No. Mahasiswa 16522070

PROGRAM STUDI TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
2021

SURAT KETERANGAN PENELITIAN

UNIVERSITAS ISLAM INDONESIA

FAKULTAS TEKNIK INDUSTRI

Gedung KH. Mas Mansur
Jl. Kaliurang Km 14,5 Yogyakarta
Telp. (0274) 895287, 898444 ext 2511;
Fax. (0274) 895007

SURAT KETERANGAN PENELITIAN

Nomor : 215/A/Ka.Lab DATMIN/FTI-UII/III/2021

Assalamu'alaikum Warahmatullahi Wabarakatuh

Kami yang bertanda tangan dibawah ini, menerangkan bahwa mahasiswa dengan keterangan sebagai berikut :

Nama	:	Muhammad Agung Pratama
No. Mhs	:	16522070
Dosen Pembimbing	:	Annisa Uswatun Khasanah, S.T., M.Sc.

Telah selesai melaksanakan penelitian yang berjudul **"Analisis Perbaikan Kualitas Layanan pada Aplikasi Zoom Cloud Meetings Berdasarkan Ulasan End-User Menggunakan Metode Text Mining Fishbone Diagram"** di Laboratorium Data Mining , Prodi Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia tercatat mulai tanggal 1 November sampai dengan tanggal 10 November 2020

Demikian surat keterangan kami keluarkan, agar dapat dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Dikeluarkan : di Yogyakarta
Tanggal : 08 Maret 2021

Mengetahui,
Kepala Lab. Data Mining

Annisa Uswatun Khasanah, S.T., M.Sc

PERNYATAAN KEASLIAN

iii

PERNYATAAN KEASLIAN

Demi Allah, saya sebagai penulis menyatakan bahwa tugas akhir ini adalah hasil kerja saya sendiri terkecuali kutipan dan ringkasan yang setiap salah satunya telah saya jelaskan sumbernya. Jika dikemudian hari ternyata terbukti pengakuan saya ini tidak benar dan melanggar peraturan yang sah dalam karya tulis dan hak kekayaan intelektual, maka saya bersedia menerima konsekuensi yang diberikan oleh Universitas Islam Indonesia dengan penuh rasa tanggung jawab.

Yogyakarta, 05 Maret 2021



(Muhammad Agung Pratama)

16522070

LEMBAR PENGESAHAN PEMBIMBING

ANALISIS PERBAIKAN KUALITAS LAYANAN PADA APLIKASI *ZOOM CLOUD MEETINGS* BERDASARKAN ULASAN *END-USER* MENGGUNAKAN METODE *TEXT MINING* DAN *FISHBONE DIAGRAM*

TUGAS AKHIR

Disusun oleh:

Nama	:	Muhammad Agung Pratama
NIM	:	16522070
Fak/Jurusan	:	FTI/Teknik Industri

Yogyakarta, 05 Maret 2021

Menyetujui,

Pembimbing Tugas Akhir

(Annisa Uswatun Khasanah S.T., M.Sc.)

LEMBAR PENGESAHAN PENGUJI

ANALISIS PERBAIKAN KUALITAS LAYANAN PADA APLIKASI *ZOOM CLOUD MEETINGS* BERDASARKAN ULASAN *END-USER* MENGGUNAKAN METODE *TEXT MINING* DAN *FISHBONE DIAGRAM*

Disusun oleh:

Nama	:	Muhammad Agung Pratama
NIM	:	16522070
Fak/Jurusan	:	FTI/Teknik Industri

Telah dipertahankan di depan sidang penguji sebagai satu syarat untuk memperoleh gelar Sarjana Strata-1 Teknik Industri

Yogyakarta, 25 Maret 2021

Tim Penguji

Annisa Uswatun Khasanah, S.T., MBA, M.Sc.	
Ketua	
Winda Nur Cahyo, S.T., M.T., Ph.D.	
Anggota I	
Harwati, S.T., M.T.	
Anggota II	

Mengetahui

Ketua Program Studi Teknik Industri

Universitas Islam Indonesia

Dr. Taufiq Immawan, S.T., M.M.

HALAMAN PERSEMBAHAN



Alhamdulillahirobbil'alamin. Puji syukur atas kehadirat Allah SWT dan atas semua nikmat yang telah diberikan. Dan tak lupa sholawat beriring salam kepada Baginda Nabi Agung Muhammad SAW. beserta keluarga dan sahabat-sahabatnya. Tugas akhir ini saya persembahkan kepada:

Bapak Jumakir dan Ibu Nursilawati selaku kedua orang tua saya yang tiada henti mendo'akan, dan memberikan semangat dan motivasi. Teruntuk adik tercinta, dan kerabat dekat yang telah memberikan dukungan dan motivasi yang begitu berarti kepada saya.

Tak lupa saya persembahkan pula kepada seluruh dosen dan karyawan Universitas Islam Indonesia yang telah mencurahkan ilmu dan bantuan selama saya kuliah di Universitas Islam Indonesia.

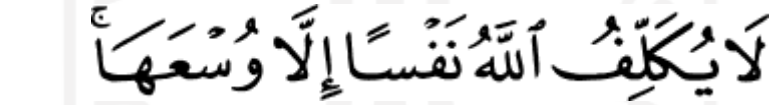
HALAMAN MOTTO


فَإِنَّ مَعَ ٱلْعُسْرِ يُسْرًا ۝٥ إِنَّ مَعَ ٱلْعُسْرِ يُسْرًا ۝٦

"*Maka sesungguhnya beserta kesulitan ada kemudahan, sesungguhnya beserta kesulitan ada kemudahan*" (Q.S. Al-Insyirah ayat 5 – 6)


يَٰٓأَيُّهَا ٱلَّذِينَ ءَامَنُواْ ٱسْتَعِينُواْ بِٱلصَّبْرِ وَٱلصَّلَوٰةِ ۚ إِنَّ ٱللَّهَ مَعَ ٱلصَّٰبِرِينَ ۝١٥٣

"*Hai orang-orang yang beriman, jadikanlah sabar dan shalat sebagai penolongmu, sesungguhnya Allah beserta orang-orang yang sabar*" (Q.S. Al-Baqarah ayat 153)


لَا يُكَلِّفُ ٱللَّهُ نَفْسًا إِلَّا وُسْعَهَا ۚ

"*Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya*" (Q.S. Al-Baqarah ayat 286)

KATA PENGANTAR



Assalamu'alaikum Warahmatullaahi Wabarakaatuh

Alhamdulillahirabbil'aalamin, segala puji dan syukur bagi Allah SWT atas berkah, rahmat, dan nikmat yang diberikan kepada hamba-Nya. Tidak lupa shalawat serta salam kepada Nabi agung Muhammad SAW beserta keluarga dan sahabat yang telah membawa umatnya dari zaman kegelapan menuju zaman terang benderang, dari zaman jahiliyah menuju zaman yang penuh dengan ilmu pengetahuan.

Atas rahmat dan izin Allah SWT, penulis dapat menyelesaikan tugas akhir dengan judul **"ANALISIS PERBAIKAN KUALITAS LAYANAN PADA APLIKASI *ZOOM CLOUD MEETINGS* BERDASARKAN ULASAN *END-USER* MENGGUNAKAN METODE *TEXT MINING* DAN *FISHBONE DIAGRAM*"** yang disusun sebagai salah satu persyaratan dalam menyelesaikan jenjang strata satu di Program Studi Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Indonesia. Dalam penelitian tugas akhir ini, penulis mendapatkan pengetahuan, bimbingan, dukungan, dan saran dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terimakasih kepada:

1. Bapak Prof. Dr. Ir. Hari Purnomo, M.T. selaku Dekan Fakultas Teknologi Industri Universitas Islam Indonesia.
2. Bapak Muhammad Ridwan Andi Purnomo, S.T., M.Sc., Ph. D. selaku Ketua Jurusan Teknik Industri Fakultas Teknologi Industri Universitas Islam Indonesia.
3. Bapak Dr. Taufiq Immawan, S.T., M.M. selaku Ketua Program Studi Teknik Industri Universitas Islam Indonesia.
4. Ibu Annisa Uswatun Khasanah, S.T., M.Sc. selaku dosen pembimbing Tugas Akhir program studi Teknik Industri Universitas Islam Indonesia.
5. Bapak Jumakir dan Ibu Nursilawati selaku orang tua saya serta adik tersayang Sophie Ananda Rizki yang selalu memberikan kasih sayang dan dukungan kepada saya hingga saat ini.

6. Teman-teman Lembaga Dakwah Kampus HAWASI. Terutama teman-teman divisi Pendidikan dan Pengembangan Sumber Daya Manusia (PPSDM) yang senantiasa memberikan dukungan kepada penulis.
7. Reza Frilian Artha, Andre Saefulloh, M. Dzaki Agung Ghifari, Adli Kurniawan, Dynta Adiasa, Dentoro Abrilicto selaku teman sehobi yang senantiasa memberikan dukungan kepada penulis.
8. Teman dekat semasa kuliah Alfian Willy Saputra, Huzein Mulyawan, Munawar Hakki, Auliya Firdaus, Hariyo Ibrahim Malik dan Abdul Gofur Wiguna yang senantiasa dengan senang hati berbagi ilmu dan terimakasih atas dukungan dan semangat yang diberikan kepada penulis.
9. Teman-teman serta sahabat Teknik Industri Fakultas Teknologi Industri Universitas Islam Indonesia terutama angkatan 2016.

Penulis berharap semoga seluruh pihak yang terlibat tetap diberikan nikmat jasmani dan rohani, serta dilancarkan segala urusannya oleh Allah SWT.

Dalam penulisan Tugas Akhir ini, penulis menyadari masih banyak kekurangan yang perlu diberikan kritik dan saran yang membangun. Oleh karena itu, penulis memohon maaf yang sebesar-besarnya kepada pihak yang terlibat serta pembaca Tugas Akhir ini. Namun terlepas dari itu semua, penulis berharap Tugas Akhir ini dapat memberikan manfaat bagi penulis dan para pembaca.

Wassalamu'alaikum warahmatullahi wabarakatuh

Yogyakarta, 05 Maret 2021



Muhammad Agung Pratama

ABSTRAK

Pada era revolusi industri 4.0 seperti saat ini, terjadi suatu perkembangan teknologi yang begitu pesat dan sudah merambah ke berbagai bidang khususnya di bidang telekomunikasi. Salah satu kemajuan teknologi telekomunikasi ialah teknologi yang dapat mendukung komunikasi *audio* maupun *video* secara bersamaan dan *real time* yaitu *video conference.* Salah satu perusahaan yang bergerak pada bidang jasa telekomunikasi berbasis *video conference* yaitu *Zoom Technologies, Inc* yang meluncurkan aplikasi *zoom cloud meetings* yang dapat diunduh melalui *Google Play Store* yang tujuannya agar pengguna *zoom* juga dapat melakukan *meeting* melalui *smartphone* mereka. Namun, hingga Desember 2020 aplikasi *Zoom Cloud Meetings* hanya mendapatkan *rating* sebesar 3,6 bintang dari total keseluruhan *rating* sebesar 5,0. Oleh karena itu, diperlukan adanya penelitian terkait *Zoom Cloud Meetings* dengan menggunakan ulasan pengguna dari *Google Play Store*. Ulasan yang dianalisis pada penelitian berjumlah sebesar 75.454 yang telah diklasifikasikan menjadi kelas sentimen positif dan negatif menggunakan algoritma *Naïve Bayes Classifier* serta dilakukan *association rule*. Klasifikasi dengan *Naïve Bayes Classifier* dengan data latih 85% dan data uji 15% menghasilkan akurasi optimum sebesar 86,67%. Dari hasil analisis ulasan, didapatkan delapan kata kunci permasalahan yaitu, "*meeting", "video", "audio", "class", "work", "update", "time"* dan "*connect"*. Beberapa penyebab permasalahan diantaranya *connecting* ke *meeting* dalam waktu yang lama, keluhan akan kualitas suara, video dan waktu yang *limited* bagi pengguna gratis, serta masalah terkait aplikasi yang tidak dapat dijalankan. Solusi yang mungkin dapat dilakukan diantaranya melakukan *update* untuk menghilangkan *bug* dan memperbaiki masalah yang dialami pengguna, meningkatkan kualitas suara dan video serta memberlakukan sistem *unlimited time* bagi semua pengguna.

Kata Kunci: *Text Mining,* Analisis Sentimen, *Naïve Bayes Classifier, Zoom Cloud Meetings,* Diagram *Fishbone, Association Rules*

DAFTAR ISI

HALAMAN JUDUL i
SURAT KETERANGAN PENELITIAN ii
PERNYATAAN KEASLIAN iii
LEMBAR PENGESAHAN PEMBIMBING iv
LEMBAR PENGESAHAN PENGUJI v
HALAMAN PERSEMBAHAN vi
HALAMAN MOTTO vii
KATA PENGANTAR viii
ABSTRAK x
DAFTAR ISI xi
DAFTAR TABEL xiv
DAFTAR GAMBAR xv
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 5
1.3 Batasan Masalah 5
1.4 Tujuan Penelitian 6
1.5 Manfaat Penulisan 6
1.6 Sistematika Penulisan Laporan 6
BAB II TINJAUAN PUSTAKA 8
2.1 Kajian Induktif 8
2.2 Kajian Deduktif 15
2.2.1 *Video Conference* 15
2.2.2 Zoom 16

2.2.3 Google Play 18
2.2.4 *Online Review* 18
2.2.5 *Data Mining* 19
2.2.6 *Web Scraping* 20
2.2.7 *Machine Learning* 21
2.2.8 *Text Mining* 22
2.2.9 *Sentiment Analysis* 23
2.2.10 *Classification* 24
2.2.11 Ukuran Evaluasi Model Klasifikasi 24
2.2.12 *Naïve Bayes Classifier* 26
2.2.13 *Association Rule* 26
2.2.14 Diagram Fishbone 27
BAB III METODOLOGI PENELITIAN 28
3.1 Rancangan Penelitian 28
3.2 Populasi dan Sampel 28
3.3 Pengumpulan Data 29
3.3.1 Metode Pengumpulan Data 29
3.4 Jenis dan Sumber Data 29
3.5 Metode Analisis Data 30
3.6 Alur Penelitian 30
BAB IV PENGUMPULAN DAN PENGOLAHAN DATA 36
4.1 Pengumpulan Data 36
4.2 Pengolahan Data 39
4.2.1 Analisis Deskriptif 39
4.2.2 *Text Preprocessing* 41
4.2.3 Representasi Model 45
4.2.4 Pembagian Data Latih dan Data Uji 46
4.2.5 Visualisasi 51

4.2.6 *Association Rule* 54
4.2.7 *Fishbone Diagram* 59
BAB V ANALISIS DAN PEMBAHASAN 64
5.1 Gambaran Umum Persepsi Pengguna *Zoom Cloud Meetings* 64
5.2 Hasil Klasifikasi Metode *Naïve Bayes Classifier* 65
5.3 Hasil *Association Rules* 65
5.4 Rencana Perbaikan Berdasarkan Diagram *Fishbone* 69
BAB VI KESIMPULAN DAN SARAN 71
6.1 Kesimpulan 71
6.2 Saran 72
DAFTAR PUSTAKA 73
LAMPIRAN 78

DAFTAR TABEL

Tabel 1.1 Jumlah Pengguna Aplikasi *Video Conference* 3
Tabel 2.1 *Confusion Matrix* 20
Tabel 2.2 Penelitian terdahulu 26
Tabel 4.1 Contoh Penghapusan Angka Pada Data 45
Tabel 4.2 Contoh Penghapusan Tanda Baca Pada Data 45
Tabel 4.3 Contoh Penghapusan *Emoji* Pada Data 46
Tabel 4.4 Contoh *Case Folding* Pada Data 46
Tabel 4.5 Contoh Penerapan *Tokenization* Pada Data 47
Tabel 4.6 Contoh Penerapan *Filtering Data* 47
Tabel 4.7 Hasil Pelabelan Kelas Sentimen berbasis Kamus *Lexicon* 48
Tabel 4.8 Perbandingan Jumlah Data pada Kelas Sentimen 49
Tabel 4.9 Pembagian Data Latih dan Data Uji 1 50
Tabel 4.10 Pembagian Data Latih dan Data Uji 2 50
Tabel 4.11 Pembagian Data Latih dan Data Uji 3 51
Tabel 4.12 Pembagian Data Latih dan Data Uji 4 51
Tabel 4.13 Hasil Klasifikasi 1 51
Tabel 4.14 Hasil Klasifikasi 2 52
Tabel 4.15 Hasil Klasifikasi 3 53
Tabel 4.16 Hasil Klasifikasi 4 54
Tabel 5.1 Rencana Perbaikan Masalah 73

.

DAFTAR GAMBAR

Gambar 1.1 Jumlah Pengguna Internet di Indonesia 1
Gambar 2.1 *Logo Zoom* 10
Gambar 2.2 Tahapan proses KDD 13
Gambar 3.1 *Interface appfollow.id* 32
Gambar 3.2 *Flowchart* penelitian 35
Gambar 4.1 *Interface* Beranda Appfollow 36
Gambar 4.2 Pemilihan Aplikasi 37
Gambar 4.3 *Dashboard Rating* & Ulasan 37
Gambar 4.4 *Interface Reply to Reviews* 38
Gambar 4.5 *Export Data* 38
Gambar 4.6 Hasil *Scraping* dari AppFollow 39
Gambar 4.7 Jumlah Ulasan Pengguna 40
Gambar 4 8 Persentase *Rating* 41
Gambar 4.9 *Bar Plot* Ulasan Positif 52
Gambar 4.10 *Bar Plot* Ulasan Negatif 53
Gambar 4.11 Asosiasi Kata Ulasan Positif 54
Gambar 4.12 Asosiasi Kata Ulasan Negatif 57
Gambar 4.13 Diagram Fishbone Untuk Permasalahan *Update* 60
Gambar 4.14 Diagram Fishbone Untuk Permasalahan Durasi 60
Gambar 4.15 Diagram Fishbone Untuk Permasalahan Audio 61
Gambar 4.16 Diagram Fishbone Untuk Permasalahan Koneksi 61
Gambar 4.17 Diagram Fishbone Untuk Permasalahan Fitur Video 62
Gambar 4.18 Diagram Fishbone Untuk Permasalahan *Meeting* 62
Gambar 4.19 Diagram Fishbone Untuk Permasalahan Sistem 63

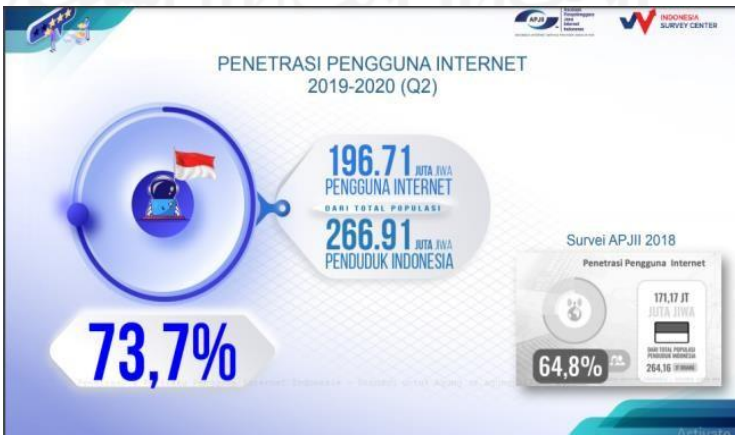
BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada era revolusi industri 4.0 seperti saat ini, terjadi suatu perkembangan teknologi yang begitu pesat dan sudah merambah ke berbagai bidang khususnya di bidang telekomunikasi. Telekomunikasi saat ini bukan lagi menjadi kebutuhan sekunder melainkan telah menjadi kebutuhan primer bagi masyarakat. Kemajuan ilmu pengetahuan dan teknologi serta kebutuhan manusia akan kemajuan telekomunikasi menyebabkan perkembangan telekomunikasi yang begitu cepat.

Perkembangan telekomunikasi ini diiringi dengan banyaknya pengguna Internet pada zaman serba digital seperti saat ini, hal ini dibuktikan dengan survei yang telah dilakukan lembaga APJII pada 2 – 25 Juni 2020 menunjukkan bahwa jumlah penetrasi pengguna Internet di Indonesia periode 2019-2020 (Q2) yaitu sebesar 73,7 %. Jumlah penetrasi tersebut meningkat sebesar 8,9% dibandingkan jumlah pengguna pada tahun sebelumnya. Terjadi kenaikan pengguna sebanyak 25 juta pengguna, yang sebelumnya sebanyak 171,7 juta menjadi 196,71 juta pengguna (Ditunjukkan pada gambar 1.1). Sampel yang digunakan pada survei ditujukan ke 7000 responden di seluruh Indonesia.



Gambar 1.1 Jumlah Pengguna Internet di Indonesia

Sumber: APJII (2020)

Seiring dengan berkembangnya zaman teknologi dan kebutuhan manusia, telekomunikasi telah berevolusi yang awalnya hanya dapat berkomunikasi melalui komunikasi suara dan teks hingga berkembang menjadi komunikasi melalui *video* secara *real time*, perkembangan ini seakan menyebabkan timbulnya tuntutan akan layanan teknologi komunikasi yang mendekati kesempurnaan, baik itu dari segi kecepatan akses, kejernihan suara, ketajaman gambar, keamanan data, dan lain sebagainya.

Salah satu teknologi modern yang dapat mendukung komunikasi *audio* maupun *video* secara bersamaan yaitu *video conference. Video Conference* merupakan salah satu jenis layanan multimedia yang dapat memenuhi kebutuhan *user* saat ini yang menginginkan adanya komunikasi secara *real time, anytime*, dan *anywhere*. Layanan ini dapat diimplementasikan melalui suatu jaringan (*network*) (Novienstinawati, 2008).

Sejak munculnya masa pandemi COVID-19, penggunaan aplikasi yang berbasis *video conference* mulai menjadi tren masyarakat karena pandemi COVID-19 yang menyebar hampir di semua negara termasuk Indonesia. Pemerintah menetapkan kebijakan baru guna menekan penyebaran virus corona dengan menghimbau masyarakat agar tidak keluar dari rumah dan bekerja dari rumah atau *Work From Home* (WFH). Begitu juga bagi pelajar dan mahasiswa untuk belajar dari rumah atau *study from home* (SFH) sehingga secara tidak langsung mewajibkan semua aktivitas tersebut harus dilakukan secara virtual atau melalui *online*. Pemberlakuan sistem tersebut mengharuskan masyarakat menggunakan aplikasi berbasis *video conference* untuk melakukan aktivitas secara daring dari rumah. Salah satu diantara perusahaan jasa layanan komunikasi *video conference* yang saat ini sedang naik daun dan banyak digunakan sejak masa pandemi COVID-19 adalah *Zoom Technologies Inc*.

Berdasarkan data dari *Statqo Analytic*s terkait dengan jumlah penggunaan aplikasi *video conference* meningkat setiap minggunya khususnya sejak dikeluarkannya kebijakan baru untuk melakukan *physical distancing* dan aktivitas *Work From Home* (WFH). Adapun pada bulan Maret 2020, jumlah penggunaan aplikasi *Zoom cloud meetings* mengalami peningkatan yang sangat signifikan. Jumlah penggunaan aplikasi *Zoom cloud meetings* meningkat hingga sekitar 183% dari 6 Maret – 26 Maret 2020 seperti yang dapat diperhatikan pada Tabel 1.1 berikut :

Tabel 1.1 Jumlah Pengguna Aplikasi *Video Conference*

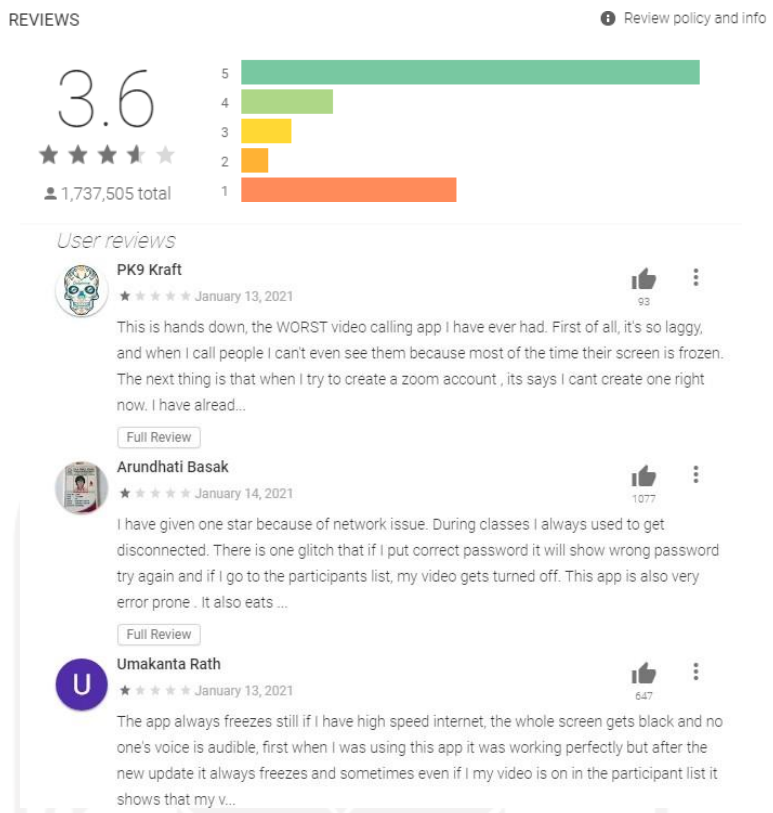
APLIKASI	JUMLAH PENGGUNA			
	28 Feb-5 Maret	6-12 Maret	13-19 Maret	20-26 Maret
Zoom Meeting	8.714	8.985	91.030	257.853
Google Meet	1.448	1.554	7.917	10.454
Skype	60.614	60.641	65.875	17.115
Cisco Webex	3.983	4.123	8.257	8.748
GoToMeeting	479	505	696	977

Sumber : (*Statqo Analytics,* 2020)

Zoom sendiri dapat diakses dengan mudah melalui website *zoom.us* dan dapat diinstal di *smartphone,* PC maupun laptop. *Zoom* juga memiliki aplikasi yang bernama *Zoom cloud meetings* yang dapat diunduh di *google play store* bagi pengguna *smartphone, zoom cloud meetings* merupakan produk utama yang diciptakan oleh *Zoom Technologies Inc*. dan aplikasi ini secara resmi dirilis pada 25 Januari 2013. Fitur utama pada aplikasi *Zoom Cloud Meetings* adalah *video conferencing* yang berbasis *cloud computing* dengan kapasitas lebih dari 100 partisipan untuk bertatap muka secara virtual. Aplikasi *zoom cloud meetings* sangat diminati oleh banyak orang karena sangat *reliable*, simpel, dengan fitur *collaboration,* webinar, presentasi dan dapat diakses di PC maupun di *smartphone*. *Zoom cloud meetings* saat ini telah di*download* lebih dari 500 juta kali di *google play. Zoom cloud meetings* telah mendapatkan ulasan dari penggunanya melalui *google play* sebanyak lebih dari 300 ribu komentar. Ulasan ini meliputi keluhan-keluhan pengguna maupun apresiasi dari pengguna.

Ulasan dari pengguna aplikasi sangat membantu perusahaan dalam memperoleh informasi terhadap citra perusahaan ataupun produk yang dibuat perusahaan itu sendiri. Hal ini disebabkan karena calon *user* umumnya mengamati perspektif dari opini pengguna lain sebelum menggunakan produk atau layanan tersebut (Fanani, 2017).

Zoom Cloud Meetings menawarkan banyak fitur dan kemudahan, akan tetapi kelebihan yang ada tersebut tidak sepenuhnya membuat pengguna merasa puas akan layanan yang telah diberikan. Berdasarkan *rating* yang didapatkan aplikasi *zoom cloud meetings* di *google play*, hingga Desember 2020 *zoom cloud meetings* hanya mendapatkan *rating* sebesar 3,6 bintang dari total penilaian sebesar 5,0 (ditunjukkan pada Gambar 1.1). Terkadang sebuah sistem juga mempunyai kelemahan, untuk itu perlu adanya perbaikan–perbaikan yang perlu dilakukan pada aplikasi *Zoom Cloud Meetings*. Kelemahan pada aplikasi *zoom cloud meetings* dibuktikan dengan keluhan-keluhan pengguna yang disampaikan melalui *user reviews* pada *google play*.



Gambar 1.2 *Rating* dan Ulasan *Zoom Cloud Meetings*

Sumber: *Google Play Store* (2020)

Berdasarkan ulasan dari pengguna tersebut dapat dianalisa faktor apa saja yang perlu diperbaiki dan dievaluasi oleh perusahaan. Untuk mengumpulkan dan mengolah semua ulasan tersebut membutuhkan waktu yang banyak jika dilakukan secara manual, maka dari itu diperlukan suatu metode atau teknik untuk dapat melakukannya. Salah satu metode yang sesuai yaitu dengan menggunakan *text mining*.

Text mining merupakan sebuah proses pengambilan informasi dari data tekstual yang memiliki kualitas tinggi serta dapat mengetahui permasalahan dalam teks dari sebuah topik tertentu. *Text mining* dalam analisa sentimen mampu mengidentifikasi emosional terhadap suatu pernyataan (Zhang, et al, 2015).

Dari latar belakang yang telah diuraikan tersebut, perlu dilakukan analisis lebih lanjut mengenai ulasan pengguna *Zoom Cloud Meetings* di *Google Play*. Pada tahap awal penelitian dilakukan pengambilan data dengan teknik *scraping* data karena jika dilakukan secara manual membutuhkan waktu untuk jumlah ulasan yang begitu banyak. Tahap selanjutnya dilakukan pengklasifikasian ulasan menjadi ulasan positif dan negatif dengan menggunakan algoritma *Naïve Bayes Classifier*. *Naïve Bayes* memiliki keunggulan dibandingkan metode klasifikasi lainnya, *naïve bayes* merupakan metode paling sederhana dan memiliki tingkat akurasi yang sangat tinggi ketika diterapkan pada

big data (Wati, 2016). Selain itu *naïve bayes* memiliki tingkat kesalahan yang sangat minimum dibandingkan algoritma lainnya (Liu et al., 2013).

Setelah tahapan klasifikasi, informasi pada ulasan dengan sentimen positif dan negatif dilakukan ekstraksi dan eksplorasi pada tahap analisis deskriptif, kemudian pada tahapan *association rules* bertujuan untuk mencari kata kunci yang sering muncul dan menemukan suatu korelasi kata (Dunham, 2003). Setelah itu dilakukan analisis diagram *fishbone* untuk menemukan akar permasalahan pada kata kunci yang telah didapatkan dari ulasan negatif yang kemudian dapat dijadikan referensi untuk rencana perbaikan terhadap permasalahan yang terdapat pada aplikasi *Zoom Cloud Meetings*. Harapannya, penelitian ini mampu memberikan informasi ataupun rekomendasi dalam suatu pemecahan masalah serta penyajian informasi dari data yang diamati dapat memberikan informasi yang berguna bagi pihak yang membutuhkan.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan, maka dalam penelitian ini permasalahan yang akan dirumuskan diantaranya adalah:

1. Bagaimana gambaran umum tentang ulasan pengguna *Zoom Cloud Meetings* berdasarkan situs *Google Play* ?
2. Berapa akurasi algoritma *Naïve Bayes* dalam mengklasifikasikan ulasan pengguna *Zoom* menjadi kelas ulasan positif dan negatif?
3. Berdasarkan *word cloud* dan *association rule*, pola kata apa saja yang sering muncul dan informasi apa yang didapatkan mengenai ulasan pengguna *Zoom Cloud Meetings* berbahasa Inggris berdasarkan situs *Google Play*?
4. Berdasarkan diagram *fishbone*, faktor-faktor apa saja yang harus dilakukan untuk memperbaiki dari hasil ulasan negatif yang didapat?

1.3 Batasan Masalah

Batasan masalah yang ditentukan untuk menghindari perluasan pembahasan dalam penelitian ini adalah sebagai berikut:

1. Data yang akan digunakan hanya data ulasan pengguna *Zoom Cloud Meetings* pada *Google Play*.
2. Penelitian ini menggunakan data ulasan pengguna aplikasi *Zoom Cloud Meetings* yang berbahasa Inggris yang tercatat dari 14 Maret 2019-20 November 2020.

3. Penelitian ini menggunakan bantuan *software python* versi 3.8.6 dan *Microsoft excel* untuk melakukan tahapan proses analisis data serta aplikasi *appfollow* untuk melakukan *scraping* data di *google play*.
4. Penelitian ini yang dilakukan dengan metode klasifikasi tidak sampai ke tahap memprediksi data yang baru.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut :

1. Mengetahui gambaran umum data ulasan tentang *Zoom Cloud Meetings* berdasarkan situs *Google Play*.
2. Mengukur akurasi algoritma *Naïve Bayes* dalam mengklasifikasikan ulasan pengguna *Zoom* menjadi kelas ulasan positif dan negatif
3. Menemukan pola kata yang sering muncul dan mendapatkan informasi penting dari ulasan pengguna aplikasi *zoom cloud meetings* berdasarkan hasil *word cloud* dan *association rule*.
4. Mendapatkan informasi faktor apa saja yang harus diperbaiki dan rencana pemecahan masalah berdasarkan acuan diagram *fishbone*.

1.5 Manfaat Penulisan

Manfaat dari penelitian ini adalah sebagai berikut:

1. Mengetahui gambaran umum tentang ulasan pengguna aplikasi *Zoom Cloud Meetings* berdasarkan *Google Play Store*.
2. Pengelompokkan ulasan tentang *Zoom Cloud Meetings* dapat memudahkan pihak perusahaan dalam mengetahui keluhan pengguna aplikasi, sehingga dapat dijadikan sebagai acuan dalam upaya menjaga kualitas dan mengembangkan kualitas dengan cara memperbaiki kekurangan yang terdapat pada aplikasi serta evaluasi ke arah yang lebih baik.

1.6 Sistematika Penulisan Laporan

Penulisan tugas akhir ini akan disusun dalam beberapa bab yang akan dijelaskan sebagai berikut :

BAB I PENDAHULUAN

Membahas mengenai latar belakang permasalahan yang menjadi alasan penelitian ini dilakukan, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Berisi pemaparan terkait istilah-istilah yang digunakan dalam penelitian ini. Bab ini juga memaparkan penelitian-penelitian terdahulu yang berhubungan dengan permasalahan yang diteliti dan menjadi acuan konseptual.

BAB III METODOLOGI PENELITIAN

Mengandung uraian terkait rancangan penelitian yang dilakukan, populasi dan sampel penelitian, metode pengumpulan data, jenis dan sumber data, metode analisis data serta diagram alur penelitian.

BAB IV PENGUMPULAN DAN PENGOLAHAN DATA

Memuat mengenai data yang diperoleh selama penelitian dan bagaimana pengolahan data tersebut. Hasil pengolahan data menjadi acuan untuk pembahasan yang akan ditulis pada sub bab pembahasan hasil.

BAB V ANALISIS DAN PEMBAHASAN

Membahas mengenai analisa yang dilakukan terhadap hasil pengumpulan, pengolahan dan analisa data yang diperoleh dari hasil penelitian sehingga menghasilkan sebuah saran rekomendasi perbaikan.

BAB VI PENUTUP

Berisi mengenai kesimpulan yang diperoleh dari hasil penelitian dan analisa data yang telah dilakukan serta saran-saran yang dapat diterapkan dari hasil pengolahan data yang dapat menjadi masukan yang berguna kedepannya.

DAFTAR PUSTAKA

LAMPIRAN

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Induktif

Adiyana et al., (2015) dalam penelitiannya menggunakan metode *text mining* dan *word cloud* bertujuan untuk menganalisis topik – topik terkait "Kpk dan Jokowi". Penelitian ini dilakukan dengan bantuan *software "R*". Hasil penelitian menunjukkan bahwa terdapat kesamaan topik utama yang digunakan pada *tweet-tweet* ketika pengguna *twitter* mencari informasi seputar topik KPK dan Jokowi pada *search engine twitter* yaitu topik utama KPK dan Polri. Berdasarkan ukuran asosiasi kata dengan nilai korelasi tidak kurang dari 0.30, kata-kata yang berasosiasi dengan kata KPK adalah kata Polri dan Lapor. Sedangkan kata-kata yang berasosiasi dengan Jokowi dimana nilai korelasi tidak kurang dari 0.30 adalah kata Widodo, Menghadiri, Izin, Pintu, Satu, Investor, Urus, Presiden, Nilai, Aktif, Bahaya, Manuver, Menang, Mulai, Relawan, dan Sejumlah. Terdapat kesamaan topik yang dibahas pada pengelompokan *tweet* untuk topik KPK dan Jokowi yaitu *tweet* yang membahas mengenai topik "kpk, polri, dan jokowi".

Ma'arif (2016) melakukan penelitian mengenai integrasi laman web tentang pariwisata Daerah Istimewa Yogyakarta yang memanfaatkan teknologi *web scraping* dan *text mining*, penelitian ini bertujuan untuk eksplorasi laman-laman yang menyajikan informasi mengenai pariwisata DIY kemudian mengumpulkan informasi tersebut secara otomatis dengan teknik *web scraping*, informasi yang telah terkumpul, selanjutnya diolah dengan metode *text mining*. Dari lima situs yang diambil untuk diolah, didapatkan hasil berupa enam kategori dan kata kunci untuk setiap kategorinya, kategori

yang didapatkan ialah pantai, alam, belanja, kuliner, budaya dan taman. Hasil akhir dari penelitian ini membangun sebuah laman web baru yang menyajikan informasi-informasi yang sudah diperoleh dengan kategorisasi yang sudah didapatkan.

Karyadi et al., (2016) melakukan penelitian terkait penggunaan metode *text mining* untuk mengetahui kecenderungan topik pemberitaan dan mengetahui topik apa yang sering muncul. Analisis data pada penelitian ini menggunakan metode *text mining* dengan bantuan *software* R dan metode *k-means clustering*. Hasil penelitian berdasarkan *word cloud* yang telah dibuat menunjukkan bahwa kata yang paling sering muncul adalah "ahok" dengan kemunculan terbanyak yaitu lebih dari 25 kali. Selain mengenai ahok kecenderungan pemberitaan lainnya yakni mengenai Rp (rupiah), DKI, mobil, macet, Luhut, RI, DPR, Jokowi, Eropa dan lain-lain. *Cluster tweet* yang terbentuk dari *K-means clustering* yakni sebanyak tiga *cluster*.

Priyanto et al., (2018) melakukan sebuah penelitian mengenai implementasi *web scraping* dan *text mining* untuk akuisisi dan kategorisasi informasi laman web tentang hidroponik. Penelitian ini bertujuan untuk mengimplementasikan teknik *web scraping* yang dikombinasikan dengan *text mining* untuk secara otomatis mengakuisisi informasi dari laman-laman web yang memuat informasi mengenai hidroponik dan mengkategorisasikannya. Hasil penelitian ini adalah data mengenai artikel hidroponik yang tersebar di berbagai laman web dapat diakuisisi secara otomatis kemudian dikelompokkan menjadi 12 topik menggunakan algoritma *topic modelling* yakni *Latent Dirichelet Allocation* (LDA).

Kurniasari (2018) dalam penelitiannya menerapkan teknik *web scraping* dan klasifikasi sentimen dengan menggunakan metode *Support Vector Machine* (SVM). Klasifikasi kemudian dianalisis dengan menggunakan *text mining*. Klasifikasi dengan metode SVM menunjukkan tingkat akurasi sebesar 96,07% pada ulasan berbahasa Inggris dan sebesar 84,07% pada ulasan berbahasa Indonesia. Berdasarkan diagram *fishbone*, hasil yang didapatkan adalah urutan faktor yang harus dibenahi menurut hasil dari ulasan negatif adalah faktor fasilitas, pelayanan dan lingkungan.

Fikria (2018) dalam penelitiannya melakukan analisis klasifikasi sentimen pada ulasan aplikasi *e-ticketing*. Dengan klasifikasi ulasan pengguna berdasarkan jenis sentimen akan mempermudah PT. KAI dalam mendapatkan informasi persepsi dari pengguna aplikasi KAI Access. Penelitian ini menerapkan tahap *preprocessing* untuk mengolah data mentah suatu ulasan agar dapat dibuat suatu model yang dapat

mempresentasikan suatu data, data ini diolah dengan menggunakan metode *Support Vector Machine* (SVM). Klasifikasi dengan algoritma SVM menghasilkan akurasi klasifikasi sebesar 89,06% pada ulasan KAI Access dan 84,68% pada ulasan Tiket.com. Kemudian eksplorasi data ulasan menggunakan metode statistika deskriptif dan asosiasi teks untuk mencari kata-kata yang berhubungan.

Titania (2020) dalam penelitiannya menerapkan *text mining* dan *social network analysis* pada jejaring sosial *twitter*, dibuatnya penelitian ini bertujuan untuk mengetahui tanggapan dari masyarakat dan mengidentifikasi aktor penting dalam diskusi kasus dugaan korupsi Asuransi Jiwasraya dan ASABRI di jejaring sosial, maka dari itu penelitian ini menggunakan metode *text mining* dan *social network analysis*. Berdasarkan hasil penelitian yang telah didapatkan ditunjukkan bahwa kata yang paling sering muncul pada data kasus dugaan korupsi Asuransi Jiwasraya adalah "pansus", "korupsi", dan "uang", sedangkan pada kasus dugaan korupsi ASABRI kata yang sering digunakan ialah "jiwasraya", "rampok", dan "pilpres". Sedangkan untuk aktor yang terpenting pada kasus dugaan korupsi Asuransi Jiwasraya dan kasus dugaan korupsi ASABRI secara berturut-turut adalah Jansen Sitindaon dan Muhammad Said Didu.

Chintalapudi et al., (2020) dalam penelitiannya melakukan analisis sentimen dan *text mining* pada dokumen medis pelaut, hal ini didasari karena pelaut lebih rentan mengalami kecelakaan dan rentan terhadap kesehatan karena budaya kerja, perubahan iklim dan kebiasaan pribadi maka dari itu dengan penerapan *text mining* dapat mengetahui lebih dalam mengenai masalah medis yang sering terjadi di kapal. Data yang akan digunakan untuk analisis adalah data pasien selama periode 2018-2020 yang dikumpulkan dari sistem kesehatan *digital Centro Internazionale Radio Medico* (C.I.R.M.). Penelitian ini menggunakan algoritma *Naïve Bayes* dan *lexicon* dan eksperimen menggunakan *tools software* R. Hasil yang didapatkan menunjukkan bahwa berdasarkan *text mining* dan *document term matrix* pembengkakan, nyeri, dan masalah perut yang memiliki korelasi paling tinggi minimal 96%. Berdasarkan *word cloud*, kata yang sering muncul adalah nyeri, bengkak, pendarahan, luka, luka, penyakit kulit, demam, sakit perut, dll, hal ini menunjukkan bahwa kecelakaan di tempat kerja paling umum terjadi diakibatkan kesalahan manusia (*Human Error*) walaupun sudah mentaati SOP yang ada. Dalam penelitian ini juga didapatkan persentase korelasi antar gejala dan hasil diagnosis mencapai 96% . Analisis sentimen divalidasi dengan *Naïve Bayes* dan menghasilkan akurasi, presisi, dan *recall* lebih dari 80%.

Pradana (2020) melakukan penelitian terkait penggunaan fitur *word cloud* dan *document term matrix* dalam *text mining*, pada penelitiannya dilakukan perbandingan kata yang sering muncul terhadap *platform facebook* dan *twitter* mengenai topik maskapai penerbangan. Penelitian ini menggunakan metode *text mining*, *word cloud* dan *term document matrix* serta menggunakan aplikasi R. Hasil berdasarkan *word cloud* dan *term document matrix* menunjukkan bahwa kata yang sering muncul pada *platform facebook* adalah "garuda" dengan frekuensi sebesar 3621 kata, sedangkan pada *Twitter* adalah kata "Indonesia" dengan frekuensi sebesar 1572 kata. *Facebook* memiliki kata-kata yang lebih cenderung positif, sedangkan *Twitter* lebih cenderung ke arah sentimen negatif.

Carracedo et al., (2020) dalam penelitian yang dilakukannya terkait analisis *text mining* untuk mengidentifikasi jalur penelitian saat ini yang dikembangkan seputar COVID-19 dan dampaknya terhadap lingkungan bisnis. Analisis pada penelitian ini menggunakan *tools software* R dan terdapat 16 *paper* yang akan dianalisis. Hasil penelitian yang didapatkan dari hasil *text mining* dan *document term matrix* menunjukkan bahwa didapatkan tiga *cluster* dengan masing-masing kata kunci. *Cluster* 1 berkaitan dengan dampak pada model manajemen kesehatan publik, *cluster* 2 membahas mengenai perubahan kebiasaan konsumsi masyarakat, dan *cluster* 3 berkaitan dengan dampak ekonomi pada organisasi bisnis saat pandemi COVID-19.

Kajian induktif berisi tentang penelitian-penelitian sebelumnya yang berkaitan dengan penelitian yang telah dilakukan. Berdasarkan penelitian terdahulu dapat diketahui beberapa kesimpulan mengenai *web scraping*, *text mining*, analisis sentimen dan asosiasi dalam tabel berikut:

Tabel 2.1 Penelitian terdahulu

Peneliti	**Metode**	**Hasil Penelitian**
(Adiyana et al., 2015)	*Text Mining,Word Cloud* dan Asosiasi Kata	Kata-kata yang berasosiasi dengan kata KPK adalah kata Polri dan Lapor. Sedangkan kata-kata yang berasosiasi dengan Jokowi adalah kata Widodo, Menghadiri, Izin, Pintu, Satu, Investor, Urus, Presiden, Nilai, Aktif,

Peneliti	Metode	Hasil Penelitian
		Bahaya, Manuver, Menang, Mulai, Relawan, dan Sejumlah. Diantara dua topik tersebut memiliki kesamaan *tweet* yaitu topik "kpk, polri, dan jokowi"
(Ma'arif, 2016)	*Web Scraping* dan *Text Mining*	Dari lima situs yang diambil untuk diolah, didapatkan hasil berupa enam kategori dan kata kunci untuk setiap kategorinya, kategori yang didapatkan ialah pantai, alam, belanja, kuliner, budaya dan taman.
(Karyadi et al., 2016)	*Text Mining, Word Cloud* dan *K-Means Clustering*	Kata yang paling sering muncul adalah "ahok" dengan kemunculan terbanyak yaitu lebih dari 25 kali. Kecenderungan pemberitaan lainnya yakni mengenai Rp (rupiah), DKI, mobil, macet, Luhut, RI, DPR, Jokowi, Eropa dan lain-lain. *Cluster tweet* yang terbentuk dari *K-means clustering* yakni sebanyak tiga *cluster*.
(Priyanto et al., 2018)	*Web Scraping* dan *Text Mining*	Data mengenai artikel hidroponik dapat diakuisisi secara otomatis kemudian dikelompokkan ke dalam 12 topik dengan klaster dan kata kunci yang sesuai.
(Kurniasari, 2018)	*Web Scraping,* SVM, dan *Fishbone Diagram*	Klasifikasi dengan metode SVM menunjukkan tingkat akurasi sebesar 96,07% pada ulasan berbahasa Inggris dan sebesar 84,07% pada ulasan berbahasa Indonesia. Berdasarkan diagram *fishbone*, hasil yang

Peneliti	Metode	Hasil Penelitian
		didapatkan adalah urutan faktor yang harus dibenahi menurut hasil dari ulasan negatif adalah faktor fasilitas, pelayanan dan lingkungan.
(Fikria, 2018)	SVM dan asosiasi teks	Klasifikasi dengan algoritma SVM menghasilkan akurasi klasifikasi sebesar 89,06% pada ulasan KAI *Access* dan 84,68% pada ulasan Tiket.com.
(Titania, 2020)	*Text Mining* dan *Social Network Analysis*	Kata yang paling sering muncul pada data kasus dugaan korupsi Asuransi Jiwasraya adalah "pansus", "korupsi", dan "uang", sedangkan pada kasus dugaan korupsi ASABRI kata yang sering digunakan ialah "jiwasraya", "rampok", dan "pilpres". Aktor yang penting dalam diskusi ini secara berturut-turut adalah Jansen Sitindaon dan Muhammad Said Didu.
(Chintalapudi et al., 2020)	*Text Mining,Document Term Matrix, Naïve Bayes* dan *Word Cloud*	Berdasarkan *text mining* dan *document term matrix* pembengkakan, nyeri, dan masalah perut yang memiliki korelasi paling tinggi minimal 96%. Berdasarkan *word cloud*, kata yang sering muncul adalah nyeri, bengkak, pendarahan, luka, luka, penyakit kulit, demam, sakit perut, dll. Analisis

Peneliti	Metode	Hasil Penelitian
		sentimen divalidasi dengan *Naïve Bayes* dan menghasilkan akurasi, presisi, dan *recall* lebih dari 80%.
(Pradana, 2020)	*Text Mining, Word Cloud* dan *Term Document Matrix*	Kata "garuda" merupakan kata yang paling sering muncul pada *platform facebook* dengan frekuensi sebesar 3621 kata, sedangkan pada Twitter adalah kata "Indonesia" dengan frekuensi sebesar 1572 kata. *Facebook* memiliki kata-kata yang lebih cenderung positif, sedangkan *Twitter* lebih cenderung ke arah sentimen negatif.
(Carracedo et al., 2020)	*Text Mining, Document Term Matrix*	Didapatkan tiga *cluster* dengan masing-masing kata kunci. *Cluster* 1 berkaitan dengan dampak pada model manajemen kesehatan publik, *cluster* 2 membahas mengenai perubahan kebiasaan konsumsi masyarakat, dan *cluster* 3 berkaitan dengan dampak ekonomi pada organisasi bisnis saat COVID-19.

Berdasarkan kajian empiris, diketahui bahwa *text mining* dapat diaplikasikan ke bidang politik, kesehatan, bisnis dan lain sebagainya. Pada penelitian terdahulu diketahui bahwa metode yang digunakan seperti *Support Vector Machine, Naïve Bayes* & *Term Document Matrix*. Namun pada penelitian ini lebih membahas terkait *text mining* pada

bidang jasa telekomunikasi dan penelitian ini menggunakan metode klasifikasi dengan algoritma *Naïve Bayes, confusion matrix* & *fishbone diagram*. Berdasarkan penelitian terdahulu diketahui bahwa metode *Naïve Bayes* merupakan metode paling sederhana dari pengklasifikasian probabilitas, memiliki tingkat akurasi yang sangat tinggi ketika diaplikasikan pada *database* dengan jumlah data yang besar, maka dari itu dalam penelitian ini menggunakan metode *Naïve Bayes*. Pada sebagian penelitian terdahulu metode yang digunakan lebih berfokus ke hasil akhir klasifikasi sedangkan penelitian ini menganalisis sampai ke akar permasalahan suatu aplikasi dan memberikan rencana perbaikan dari permasalahan tersebut. Penelitian ini diharapkan dapat menjadikan gambaran dan bahan evaluasi kinerja dari perusahaan di bidang jasa telekomunikasi, khususnya aplikasi berbasis *video conference* seperti *zoom cloud meetings*.

2.2 Kajian Deduktif

2.2.1 *Video Conference*

Video conference merupakan suatu telekomunikasi berupa *audio* dan *video* untuk menghubungkan orang ke tempat berbeda dalam waktu yang bersamaan untuk pertemuan (Herryawan, 2009). *Video conference* adalah salah satu sarana yang memungkinkan komunikasi data, suara, dan gambar yang bersifat *duplex* serta *real time* (Nurdiansyah, 2013).

Menurut (Gough, 2006), *video conferencing* dapat dibagi menjadi 3 jenis, yaitu :

1. *Personal video conferencing*

   Dalam *personal video conferencing*, terdapat komunikasi yang berupa *video* dan *audio* antara dua orang yang berinteraksi. Dapat ditambahkan fitur tambahan berupa pengiriman teks, seperti yang terdapat pada kebanyakan perangkat lunak *Instant Messaging* (IM), seperti *Windows Live Messenger* dan *Yahoo! Messenger.*

2. *Business video conferencing*

   Business video conferencing memiliki fitur yang sama dengan *personal video conferencing* ditambah dengan beberapa fitur seperti:

   - Kemampuan untuk berkomunikasi tidak hanya antara dua orang, namun bisa lebih
   - Fitur untuk berbagi file (*file sharing*)
   - Kemampuan untuk melakukan presentasi
   - Fasilitas *whiteboard* dan fitur-fitur lainnya

   Business video conferencing ini membutuhkan biaya yang lebih besar dibandingkan dengan personal *video conferencing*, sebagai akibat dari fasilitas-

fasilitas yang disediakan.

3. *Web video conferencing*

   Web video conferencing umumnya digunakan pada seminar yang menggunakan web, dimana pemirsa dapat melihat *video* yang dikirimkan oleh pembicara seminar. Oleh karena itu, *web video conferencing* merupakan komunikasi satu arah, karena pemirsa tidak dapat mengirimkan *video*nya kepada si pengirim.

Video conference menjadi kebutuhan yang mengalami *demand* yang tinggi sebagai satu- satunya alternatif yang paling aman dan terbukti dapat diandalkan khususnya pada saat masa pandemi COVID-19. Selain itu *video conference* juga mampu mengatasi masalah geografis karena setiap orang bisa terhubung tanpa harus memikirkan jarak.

Saat ini layanan *video conference* sudah semakin berkembang dan bahkan sudah merebak ke berbagai sektor kehidupan, misalnya: pendidikan, bisnis, kesehatan, dan lainnya (Herryawan, 2009). Persaingan antar perusahaan layanan yang mendukung fitur *video conferencing* sudah semakin ketat. Terbukti bahwa saat ini sudah ada banyak aplikasi untuk melakukan *video conferencing* yaitu seperti: *Zoom, Google Meet, Cisco Webex, Skype, Microsoft Teams*, dan lain sebagainya.

2.2.2 Zoom

Zoom Technologies, Inc didirikan pada 21 April 2011 oleh Eric Yuan yang saat ini merupakan CEO dari *Zoom Technologies, Inc* itu sendiri. *Zoom* memiliki visi perusahaan yaitu "Komunikasi *video* mendorong orang untuk meraih sesuatu yang lebih", dan misi "Menjadikan komunikasi *video* yang mudah dan aman". *Zoom* hanya menganut satu *value* saja yaitu "peduli", peduli disini maknanya adalah *Zoom* peduli dengan penggunanya, komunitasnya, perusahaannya, rekan tim dan diri sendiri.

Gambar 2.1 Logo *Zoom*

Sumber: *https://zoom.us* (2020)

Zoom menawarkan berbagai produk dan layanan antara lain adalah *Zoom Cloud Meetings* dan *Zoom rooms*. *Zoom Cloud Meetings* merupakan aplikasi *video conference* berbasis *cloud computing* yang sangat mudah digunakan, memiliki banyak ragam fitur

dan aplikasi ini dapat diakses di PC ataupun di *smartphone* (Herlinawati et al., 2020). Fitur utama pada aplikasi *zoom cloud meetings* adalah sebagai berikut:

a) *One-on-one meetings*: Pengguna melakukan *meeting* atau konferensi *video* dengan satu pengguna lainnya. Fitur ini dapat digunakan secara terus-menerus (*unlimited*) bahkan dengan *free plan* (tidak berlangganan layanan *Zoom*).
b) *Group Video Conferences*: Konferensi *video* dengan grup yang besar atau lebih dari satu orang. Layanan *Zoom* dapat mengakomodir pertemuan virtual hingga lebih dari 500 orang (jika pengguna menggunakan *add-on* tertentu). *Free plan* mengizinkan kamu untuk mengadakan konferensi *video* hingga 100 orang dengan batas waktu 40 menit.
c) *Screen sharing*: Konferensi *one-on-one* atau dengan grup besar dan *screen sharing* dengan peserta lain sehingga mereka dapat melihat apa yang kita lihat.

Masih ada beragam fitur lainnya dari *zoom* yang menjadi ciri khasnya tersendiri seperti: *virtual background, webinar*, dan lain sebagainya. Selain *zoom cloud meetings*, *zoom* juga memiliki layanan *zoom rooms*. *Zoom rooms* merupakan sistem ruangan berbasis perangkat lunak yang memberikan pengalaman terintegrasi untuk konferensi *audio*, berbagi layar nirkabel, dan konferensi *video*.

Zoom rooms dapat digunakan untuk peserta ruangan saja, atau peserta jarak jauh yang bergabung dari ruangan lain, dari *desktop* mereka, atau dari perangkat seluler mereka. Setiap ruangan memerlukan minimal *support device* seperti komputer *macOS*, tablet *iPad*, Android, atau *Windows* 10 yang menjalankan aplikasi *Zoom rooms*. Peralatan tambahan dapat mencakup layar TV, speaker, kamera, mikrofon, dan iPad untuk menjalankan tampilan penjadwalan *Zoom rooms*. Perangkat dapat didaftarkan di *Zoom Device Management* untuk menyediakan manajemen jarak jauh tambahan dari portal web *Zoom*.

Zoom merupakan perusahaan yang menawarkan kemudahan, kecepatan akses dan kehandalan serta peduli terhadap kebutuhan penggunanya. Hal ini dibuktikan dengan diraihnya berbagai bentuk penghargaan yang telah didapatkan oleh *zoom*. Pada tahun 2019 *zoom* memenangkan 8 penghargaan dan pada tahun 2020 *zoom* berhasil menambah penghargaannya menjadi 12 penghargaan. Salah satu penghargaan yang telah berhasil diraih oleh *zoom* adalah termasuk salah satu *Best Company Professional Development*. Selain itu juga, *zoom* berhasil menjadi 100 *brand* terbaik di dunia pada tahun 2020 yang diselenggarakan oleh *Interbrand*.

2.2.3 *Google Play*

Google Play adalah layanan konten digital milik *Google* yang melingkupi toko *online* untuk produk-produk seperti musik/lagu, buku, aplikasi, permainan, ataupun pemutar media berbasis *cloud*. Layanan ini dapat diakses baik melalui web, aplikasi android (*Play Store*) dan *Google TV*. *Google Play* mulai dikenalkan pada bulan Maret 2012 sebagai pengganti dari *Google Play* dan layanan musik *Google*.

Google Play milik *Google* saat ini telah menyediakan sekitar 700.000 aplikasi mobile menurut *AppBrain7*. Setelah beberapa bulan, mungkin ada lebih dari sepuluh ribu komentar tekstual dari aplikasi baru yang diluncurkan di *Google Play* (Liu et al, 2013). Hal ini sangat sesuai untuk penerapan analisis *text mining* terhadap komentar-komentar pengguna tentang aplikasi pada *Google Play* untuk membantu pihak perusahaan untuk menganalisis dan mengevaluasi aplikasinya.

2.2.4 *Online Review*

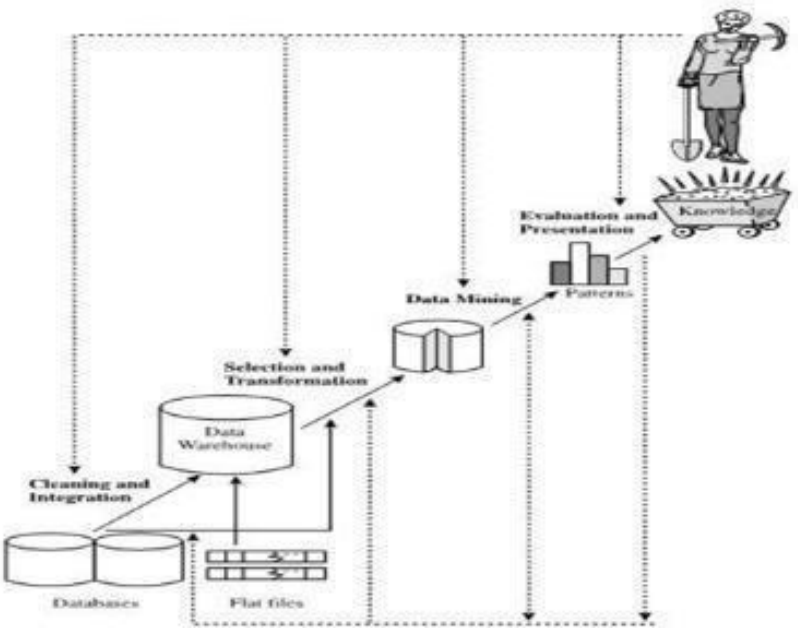
Dalam Kamus Besar Bahasa Indonesia, ulasan yang diterjemahkan ke dalam bahasa Indonesia menjadi ulasan memiliki pengertian kupasan, tafsiran, komentar, sedangkan kata *online* memiliki arti *controlled by or connected to a computer* sesuai dengan pengertian yang dijelaskan dalam *Oxford Dictionaries online* yang apabila diterjemahkan yaitu dikontrol oleh atau terkoneksi pada sebuah komputer. Menurut Mudambi dan Schuff dalam jurnal yang berjudul "*A Study of Customer Review on Amazon.com*, "*Online customer reviews can be defined as peer-generated product evaluations posted on company or third party websites*". Dalam terjemahannya yaitu *online review* dari pelanggan diartikan sebagai hasil evaluasi produk yang diunduh dalam situs perusahaan atau situs pihak ketiga.

Menurut Kotler & Amstrong (1996) ada empat metode untuk mengukur kepuasan pelanggan yang salah satunya adalah sistem keluhan dan saran melalui berbagai media seperti kolom komentar dan juga internet seperti *online review*.

2.2.5 *Data Mining*

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. Data mining adalah suatu proses yang menggunakan teknik statistik, matematika, kecerdasan tiruan, dan machine learning untuk mengekstraksi serta mengidentifikasi informasi yang bermanfaat untuk pengetahuan yang terkait dari berbagai database besar (Turban et al, 2005). Menurut Tan et al (2006) *data mining* adalah proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Istilah data mining kadang disebut juga *knowledge discovery*.

Data mining adalah bagian integral dari penemuan pengetahuan dalam database (KDD) yang merupakan proses keseluruhan mengubah data mentah menjadi pola-pola data menarik yang merupakan informasi yang dibutuhkan oleh pengguna sebagai pengetahuan. Untuk mengetahui proses *knowledge discovery* dalam *database* dapat dilihat pada Gambar 2.1 berikut :



Gambar 2.2 Tahapan proses KDD

Sumber: Han & Kamber (2006)

Han & Kamber (2006) menyatakan bahwa KDD terdiri dari langkah-langkah sebagai berikut:

1. *Data cleaning* adalah proses mengecek data yang tidak konsisten dan menghilangkan *noise*.
2. *Data integration* adalah proses pengecekan penggabungan data apabila memiliki sumber data terdiri dari banyak sumber.
3. *Data selection* adalah pemilihan data yang relevan yang akan digunakan sebelum masuk proses *data mining*.
4. *Data transformation* adalah proses transformasi data berdasarkan pada jenis atau pola informasi yang akan dicari dalam basis data. Transformasi data dapat dilakukan juga sebelum tahap proses seleksi data.
5. *Data mining* adalah suatu proses yang penting dengan melibatkan metode dan teknik tertentu untuk menemukan suatu pola data atau informasi yang menarik.
6. *Pattern evaluation* adalah proses untuk mengidentifikasi pola data yang mewakili knowledge yang ada didalam data itu sendiri.
7. *Knowledge representation* adalah proses visualisasi dan teknik menyajikan *knowledge* digunakan untuk menampilkan hasil proses data mining kepada pihak yang berkepentingan.

2.2.6 *Web Scraping*

Web Scraping adalah proses pengambilan sebuah dokumen semi-terstruktur dari internet, umumnya berupa halaman-halaman web seperti HTML atau XHTM yang kemudian dokumen tersebut dianalisis untuk diambil data tertentu dari halaman tersebut untuk digunakan bagi kepentingan lain (Turland, 2010). *Web scraping* sering dikenal sebagai *screen scraping*. *Web Scraping* tidak termasuk dalam bidang *data mining* karena *data mining* menyiratkan upaya untuk memahami pola semantik atau tren dari sejumlah besar data yang telah diperoleh (Josi et al., 2014). Pengaplikasian *web scraping* (dengan kata lain seperti: *intelligent, automated, or autonomous agents*) hanya fokus pada cara memperoleh data melalui pengambilan dan ekstraksi data dengan ukuran data yang bervariasi. *Web scraping* memiliki beberapa langkah untuk melakukannya (Juliasari & Sitompul, 2012), yaitu sebagai berikut:

1. *Create Scraping Template*

   Pembuat program mempelajari dokumen HTML dari laman web yang akan diambil informasinya untuk tag HTML yang mengapit informasi yang akan diambil.

2. *Explore Site Navigation*

   Pembuat program mempelajari teknik navigasi pada laman web yang akan diambil data informasinya untuk ditirukan pada aplikasi *web scraper* yang akan dibuat baru.
3. *Automate Navigation and Extraction*

   Berdasarkan informasi yang didapat pada langkah sebelumnya, aplikasi *web scraper* dibuat untuk mengotomatisasi pengambilan informasi dari laman web yang sudah ditentukan.
4. *Extracted Data and Package History*

   Informasi yang didapat dari langkah 3 disimpan dalam tabel atau *database tabel*.

Adapun manfaat dari pengambilan data dengan web scraping ini agar informasi yang diambil lebih terfokus sehingga memudahkan dalam melakukan pencarian sesuatu.

2.2.7 *Machine Learning*

Machine Learning (ML) atau pembelajaran mesin merupakan pendekatan dalam *Artificial Intelligence* (AI) yang banyak digunakan untuk menggantikan atau menirukan perilaku manusia untuk menyelesaikan masalah atau melakukan otomatisasi. Sesuai dengan namanya, *machine learning* mencoba menirukan bagaimana proses manusia atau makhluk cerdas belajar dan mengeneralisasi (Tanaka & Okutomi, 2014). Dalam pembelajaran *machine learning*, terdapat beberapa skenario-skenario seperti:

1. *Supervised Learning*

   Penggunaan skenario *supervised learning*, pembelajaran menggunakan masukan data pembelajaran yang telah diberi label. Setelah itu membuat prediksi dari data yang telah diberi label.
2. *Unsupervised Learning*

   Penggunaan skenario *unsupervised learning*, pembelajaran menggunakan masukan data pembelajaran yang tidak diberi label. Setelah itu mencoba untuk mengelompokan data berdasarkan karakteristik-karakteristik yang ditemui.
3. *Reinforcement Learning*

   Pada skenario *reinforcement learning* fase pembelajaran dan tes saling dicampur. Untuk mengumpulkan informasi pembelajar secara aktif dengan berinteraksi ke lingkungan sehingga untuk mendapatkan balasan untuk setiap aksi dari pembelajar.

2.2.8 *Text Mining*

Text mining adalah sebuah proses pengambilan informasi dari data tekstual yang memiliki kualitas tinggi serta dapat mengetahui permasalahan dalam teks dari sebuah topik tertentu. *Text mining* dalam analisa sentimen mampu mengidentifikasi emosional terhadap suatu pernyataan (Zhang, et al, 2015). Pekerjaan yang dilakukan dalam konsep *Text Mining* secara garis besar adalah penggalian deskriptif (*descriptive mining*) dan penggalian prediktif (*predictive mining*). Pekerjaan *predictive mining* meliputi klasifikasi dokumen ke dalam kategori-kategori, lalu menggunakan informasi tersebut untuk membuat keputusan (SAS Institute, 2010).

Perbedaan antara *text mining* dengan data mining terletak pada sumber data yang digunakan. Dalam *text mining* pola–pola yang diekstrak dari data tekstual yang tidak terstruktur bukan berasal dari suatu database. Beberapa kesamaannya adalah data yang digunakan merupakan data besar dan data berdimensi tinggi dengan struktur yang terus berubah. Dalam *data mining*, data yang diolah adalah data yang terstruktur dari proses *warehousing* sehingga lebih mudah diproses oleh mesin/komputer. Analisis teks lebih sulit karena teks umumnya berupa konsumsi manusia secara langsung bukan digunakan untuk mesin/komputer. Ditambah struktur teks yang kompleks, struktur yang tidak lengkap, bahasa yang berbeda, dan arti yang tidak standar. Oleh sebab itu pada umumnya digunakan *Natural Language Processing* untuk analisis teks yang tidak berstruktur tersebut. Tahapan-tahapan dalam *text mining* secara umum adalah *text preprocessing* dan *feature selection* (Feldman & Sanger, 2007). Dimana penjelasan dari tahap-tahap tersebut adalah sebagai berikut:

a. *Text Preprocessing*

Tahapan awal dari *text mining* adalah *text preprocessing* yang bertujuan untuk mempersiapkan teks menjadi data yang akan mengalami pengolahan pada tahapan berikutnya. Tahap ini mencakup semua rutinitas dan proses untuk mempersiapkan data yang akan digunakan pada operasi *knowledge discovery* sistem *text mining* (Feldman & Sanger, 2007). *Text preprocessing* berfungsi untuk mengubah data teks yang tidak terstruktur atau sembarang menjadi data yang terstruktur. Secara umum proses yang dilakukan dalam tahapan *preprocessing* adalah sebagai berikut:

1. *Spelling*

   Proses perbaikan kata-kata yang masih salah dalam pengejaan atau merupakan singkatan dalam bentuk tertentu. Perbaikan kata dilakukan agar jumlah perhitungan dimensi kata tidak melebar. Perhitungan jumlah dimensi kata akan melebar jika kata yang salah eja atau kata singkatan tidak dirubah karena akan dihitung sebagai kata yang berbeda pada proses penyusunan matriks.

2. *Case Folding*

   Case folding adalah mengubah semua huruf dalam dokumen menjadi satu bentuk yang sama, peran *case folding* dibutuhkan dalam mengkonversi keseluruhan teks dalam dokumen menjadi suatu bentuk standar (dalam hal ini huruf kecil atau *lowercase*).

3. *Filtering*

   Tahap filtrasi adalah tahap pengambilan kata-kata penting dengan menggunakan algoritma *stoplist*. Algoritma *stoplist* (membuang kata yang kurang penting) atau *wordlist* (menyimpan kata yang penting) dapat digunakan pada tahap ini. *Stopword* adalah kata-kata yang tidak memiliki makna dan bukan merupakan kata penting dari suatu dokumen sehingga dapat dibuang. Contoh *stopword* adalah "yang", "dan", "di", "dari" dan seterusnya (Putri, 2016). Dalam filtrasi ini menggunakan *stoplist/stopword* agar kata-kata yang kurang penting dan sering muncul dalam suatu dokumen dapat dibuang sehingga kata-kata yang penting dan memiliki makna dapat diproses ke tahap selanjutnya.

4. *Tokenization*

   Proses pemecahan teks kalimat menjadi potongan kata-kata guna membuat kalimat yang terbentuk menjadi lebih bermakna yang disebut token. Tujuan dari proses ini adalah mendapatkan potongan kata yang akan menjadi entitas yang memiliki nilai dalam matriks dokumen teks yang akan dianalisis.

b. *Feature Selection*

Kata-kata yang tidak relevan dengan proses pengkategorisasian dapat dihilangkan tanpa mempengaruhi *classifier* bahkan meningkatkan kinerja karena mengurangi *noise*. Langkah *preprocessing* ini disebut *feature selection* (Feldman & Sanger, 2007). Walaupun di tahap sebelumnya sudah dilakukan tahap penghapusan kata–kata yang tidak deskriptif (*stopwords*), tidak semua kata–kata di dalam dokumen memiliki arti penting. Maka dari itu untuk mengurangi dimensi, pemilihan hanya dilakukan pada kata–kata yang relevan dan mempresentasikan isi dari suatu dokumen. Kata-kata yang dinilai penting dilihat dari banyaknya kemunculan suatu kata dan kata yang paling informatif dari semua data

2.2.9 *Sentiment Analysis*

Analisis sentimen atau istilah lainnya *opinion mining* merupakan salah satu cabang penelitian *Text Mining*. Analisis sentimen adalah riset komputasional dari opini, sentimen dan emosi yang diekspresikan secara tekstual. Jika diberikan suatu data teks

yang berisi opini mengenai suatu objek, maka analisis sentimen bertujuan untuk mengekstrak atribut dan komponen dari objek yang telah didokumentasi pada setiap dokumen dan untuk menentukan apakah komentar tersebut mempunyai arti positif atau negatif (Rozi et al, 2012). Analisis sentimen dapat digunakan dalam berbagai macam *domain* seperti, produk konsumen, jasa kesehatan, jasa keuangan, peristiwa sosial maupun politik. Dalam dunia nyata, bisnis dan organisasi selalu ingin melihat opini publik tentang suatu produk atau jasa (Liu et al, 2012).

2.2.10 *Classification*

Klasifikasi dapat didefinisikan sebagai proses untuk menyatakan suatu objek data sebagai salah satu kategori (kelas) yang telah didefinisikan sebagai proses untuk menyatakan suatu objek data sebagai salah satu kategori (kelas) yang telah didefinisikan sebelumnya. Klasifikasi banyak digunakan dalam untuk berbagai hal, diantaranya adalah deteksi kecurangan (*fraud detection*), pengelolaan pelanggan, diagnosis medis, prediksi penjualan, dan sebagainya (Azizah, 2019).

Pada umumnya proses klasifikasi dibagi menjadi dua tahap yaitu tahap pembelajaran dan tahap pengujian. Pada tahap pembelajaran, sebagian data yang telah diketahui kelas datanya diumpankan untuk membentuk model perkiraan. Kemudian pada model tahap pengujian yang sudah terbentuk diuji dengan sebagian data lainnya untuk mengetahui akurasi dari model tersebut. Jika akurasinya mencukupi, model ini dapat dipakai untuk prediksi kelas data yang belum diketahui. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang telah diklasifikasi dan dengan menggunakan hasilnya untuk menghitung jarak antara ciri-ciri citra *template* dan citra masukan (Ulwan, 2016).

2.2.11 Ukuran Evaluasi Model Klasifikasi

Pada umumnya evaluasi terhadap suatu klasifikasi dilakukan menggunakan sebuah himpunan data uji. Model klasifikasi yang dibuat adalah pemetaan dari suatu baris data dengan keluaran sebuah hasil prediksi kelas/target dari data tersebut. Klasifikasi yang memiliki dua kelas sebagai keluarannya disebut dengan klasifikasi biner. Kedua kelas tersebut direpresentasikan dalam {0,1}, {+1,- 1} atau {*positive*; *negative*} (Kurniasari, 2018).

Dalam proses evaluasi klasifikasi terdapat empat kemungkinan hasil yang terjadi dari proses pengklasifikasian suatu baris data. Jika data positif dan diprediksi positif maka akan dihitung sebagai *true positive*, tetapi jika data itu diprediksi negatif maka akan dihitung sebagai *false negative*. Jika data negatif dan diprediksi negatif akan dihitung

sebagai *true negative*, tetapi jika data tersebut diprediksi positif maka akan dihitung sebagai *false positive* (Fawcett, 2006). Hasil klasifikasi *biner* pada suatu *dataset* dapat direpresentasikan dengan matriks 2 x 2 yang disebut *confusion matrix*.

Tabel 2.2 *Confusion Matrix*

Prediksi	Aktual	
	Positif	Negatif
Positif	*True Positive* (TP)	*False Negative* (FN)
Negatif	*False Positive* (FP)	*True Negative* (TN)

Sumber: Fawcett (2006)

Menurut Manning, Prabhakar, dan Hinrich (2009), *confusion matrix* merupakan salah satu *tools* yang digunakan pada *machine learning* untuk melakukan visualisasi dan pada umumnya memuat dua kategori atau lebih. *Confusion matrix* digunakan untuk mengevaluasi performansi model yang dibangun oleh algoritma klasifikasi. *Confusion matrix* memiliki beberapa rumus umum yang dapat digunakan untuk menghitung performa klasifikasi. Hasil dari nilai akurasi, presisi, dan *recall* dapat ditampilkan dalam persentase. Berikut rumus yang digunakan pada *confusion matrix* (Song et al, 2006):

a. Akurasi

Akurasi adalah jumlah proporsi prediksi yang benar. Adapun rumus penghitungan akurasi dapat dilihat pada persaman 2.1.

$$\text{Akurasi} = \frac{TP+FN}{TP+FP+TN+FN} \quad \dots (2.1)$$

b. Presisi

Presisi adalah proporsi jumlah dokumen teks yang relevan terkenali diantara semua dokumen teks yang terpilih oleh sistem. Rumus presisi dapat dilihat pada persamaan 2.2.

$$\text{Presisi} = \frac{TP}{TP+FP} \quad \dots (2.2)$$

c. Recall

Recall adalah proporsi jumlah dokumen teks yang relevan terkenali diantara semua dokumen teks relevan yang ada pada koleksi. Rumus *recall* dapat dilihat pada persamaan 2.3.

$$\textit{Recall} = \frac{TP}{TP+FN} \quad \dots (2.3)$$

2.2.12 *Naïve Bayes Classifier*

Naive Bayes Classifier merupakan teknik klasifikasi berdasarkan Teorema Bayes dengan asumsi independensi di antara para prediktor. *Naive Bayes Classifier* memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. *Naïve bayes* merupakan metode paling sederhana dari pengklasifikasian probabilitas, memiliki tingkat akurasi yang sangat tinggi ketika diaplikasikan pada *database* dengan jumlah data yang besar (*Big Data)* (Wati, 2016). Selain itu model *naïve bayes* juga memiliki tingkat kesalahan yang sangat minimum dibandingkan dengan algoritma klasifikasi lainnya (Liu et al., 2013).

2.2.13 *Association Rule*

Analisis asosiasi atau *association rule* adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi item (Fauzy, 2015). *Association Rule* adalah teknik *data mining* yang berguna untuk menemukan suatu korelasi atau pola yang terpenting atau menarik dari sekumpulan data besar (Dunham 2003). Motivasi awal *Association Rule* berasal dari keinginan untuk menganalisis data transaksi supermarket ditinjau dari perilaku pelanggan dalam membeli produk. *Interestingness measure* yang dapat digunakan dalam analisis asosiasi adalah (Han et al. 2012):

1. *Support*

   Nilai *support* dari aturan *A* => *B* dalam set transaksi D adalah persentase dari transaksi D yang mengandung *item A* dan *B*, seperti pada persamaan (2.1) berikut:

$$Support\ (A => B) = P\ (A \cup B) \quad \ldots. (2.1)$$

2. *Confidence*

   Nilai *confidence* dalam aturan *A* => *B* dalam set transaksi D adalah persentase dari jumlah transaksi yang mengandung A dan B, dengan jumlah transaksi yang mengandung A, seperti pada persamaan (2.2) berikut:

$$Confidence\ (A => B) = P\ (B \mid A) = \frac{\text{nilai support(A}\cup\text{B)}}{\text{nilai support A}} \quad \ldots.(2.2)$$

3. *Lift*

   Nilai *lift* adalah perhitungan korelasi sederhana. Nilai *lift* yang kurang dari 1 maka *item A* berkorelasi negatif dengan *item B*. Nilai *lift* yang lebih dari 1 menunjukkan *item A* berkorelasi positif dengan *item B* yang berarti kedua *item* saling berhubungan, seperti yang ditunjukkan pada persamaan (2.3) berikut

$$Lift\ (A, B) = \frac{P\ (A \cup B)}{P\ (A) P\ (B)} \quad \ldots. (2.3)$$

Salah satu algoritma yang digunakan untuk *association rule mining* adalah algoritma apriori. Algoritma apriori adalah suatu algoritma yang dikenal dalam melakukan pencarian *frequent itemset* dengan *association rule* (Aprianti et al. 2017). Algoritma ini menggunakan pendekatan iterasi yang dikenal sebagai "*level-wise search*", di mana *k-itemsets* digunakan untuk mendapatkan (k+1)-*itemsets*. Pertama, set dari *frequent* 1-*itemsets* ditemukan dengan pembacaan pada basis data untuk akumulasi jumlah dari masing-masing *item*, dan mengumpulkan *items* tersebut sesuai dengan *minimum support* yang ditentukan. Hasil dari set tersebut adalah L1, kemudian L1 digunakan untuk menemukan L2, dan dilakukan iterasi sampai mendapatkan Lk setelah melakukan pencarian dalam satu basis data. Cara untuk meningkatkan efisiensi dari iterasi pada Apriori digunakan sebuah properti Apriori, yang merupakan *minimum support* untuk melakukan pemangkasan *itemset* yang tidak memenuhi *minimum support* (Han et al. 2012).

2.2.14 Diagram *Fishbone*

Diagram *fishbone* dikembangkan oleh Dr. Kaoru Ishikawa pada tahun 1943, sehingga sering disebut dengan diagram Ishikawa. Diagram sebab-akibat (*cause and effect diagram* atau *fishbone diagram*) adalah sebuah teknik grafis yang digunakan untuk mengurutkan dan menghubungkan interaksi antara faktor-faktor yang berpengaruh dalam suatu proses. Diagram ini berguna untuk menganalisis dan menemukan faktor-faktor yang berpengaruh atau efek secara signifikan di dalam menentukan karakteristik kualitas *output* kerja. Efek ini dapat bernilai "baik" dan dapat bernilai "buruk". Jadi dengan diketahui sebab dari efek yang terjadi, diharapkan hasil dari proses produksi bisa diperbaiki dengan mengubah faktor terkontrol dari suatu proses. Diagram ini juga berguna untuk mengidentifikasi akar penyebab potensi dari suatu masalah. Diagram *fishbone* memfokuskan pada penekanan masalah atau gejala yang merupakan akar penyebab masalah. Diagram *fishbone* juga menampilkan penyebab-penyebab suatu masalah dengan cara menghubungkan penyebab-penyebab menjadi satu (Fauziah, 2009).

BAB III

METODOLOGI PENELITIAN

3.1 Rancangan Penelitian

Penelitian ini merupakan penelitian dengan menggunakan analisis kuantitatif untuk beberapa data yang berupa hasil *scraping data* dari ulasan *google play*. Pada proses *scraping* menggunakan aplikasi *web scraping* berbasis web yaitu, *appfollow*. Proses pengambilan data dilakukan secara online dan penggalian informasi terhadap ulasan-ulasan terhadap penggunaan aplikasi *zoom cloud meetings* dan nantinya proses *text mining* menggunakan bantuan *tool* berupa *python* 3.8.6. Setelah data primer atau data mentah telah didapatkan, dilanjutkan dengan tahap *text preprocessing* melalui tahap *cleaning, case folding, tokenizing* dan *filtering* yang bertujuan untuk membersihkan dan mempersiapkan teks menjadi data yang akan digunakan untuk tahap *word cloud*, setelah itu data diprediksi dan diklasifikasikan bantuan *tool python* 3.8.6. Data ulasan yang sudah dikelompokkan berdasarkan kategori ulasan positif dan negatif, selanjutnya dilakukan *word cloud* yang bertujuan untuk menemukan pola kata yang sering muncul terkait keluhan atau masalah yang dialami pengguna selama menggunakan aplikasi. Data pola kata dari *word cloud* dilanjutkan ke tahap asosiasi untuk menemukan hubungan antar kata baik pada ulasan positif maupun ulasan negatif dari hasil *word cloud*, sehingga mendapatkan informasi yang dapat dijadikan referensi untuk menemukan faktor-faktor apa saja yang menjadi akar permasalahan di *fishbone diagram*.

3.2 Populasi dan Sampel

Populasi dalam penelitian ini adalah *database* website *Google Play Store*, yaitu semua data *review* pengguna aplikasi *Zoom Cloud Meetings* sedangkan sampel yang digunakan dalam penelitian ini adalah *review* pengguna aplikasi *Zoom Cloud Meetings* berbahasa Inggris pada tanggal 14 Maret 2019 – 20 November 2020. Ulasan pengguna tersebut

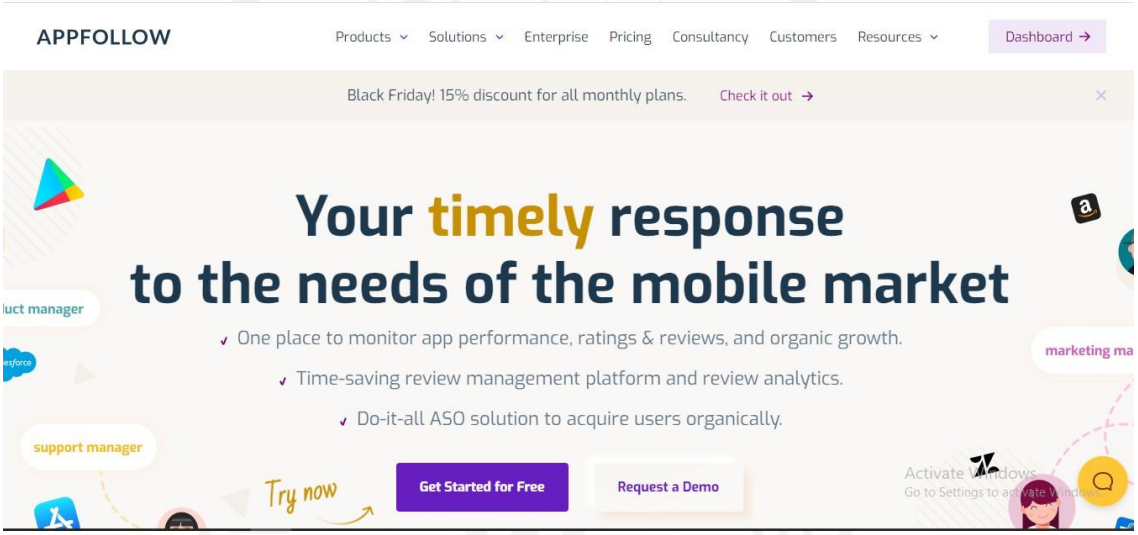
diambil dari ulasan pengguna paling awal dalam bahasa Inggris pada tanggal 14 Maret 2019 sampai hari pengambilan data yakni tanggal 20 November 2020.

3.3 Pengumpulan Data

3.3.1 Metode Pengumpulan Data

1. *Web Scraping*

   Data yang digunakan dalam penelitian yaitu berupa ulasan pengguna aplikasi *Zoom Cloud Meetings* di *Google Play Store*. Data yang dikumpulkan berupa data teks yang diambil dengan *scraping data* melalui situs web *appfollow.io*.



Gambar 3.1 *Interface appfollow.id*

Sumber: https://appfollow.io/ (2020)

2. Studi Pustaka

   Studi pustaka dilakukan untuk mencari materi dan teori yang berhubungan dengan penelitian ini dan memudahkan dalam menentukan proses yang akan dilakukan selama penelitian. Studi pustaka meliputi *text mining*, analisis sentimen, *association rule* dan *fishbone diagram*.

3.4 Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data primer. Data tersebut diperoleh dengan *scraping* dari halaman situs web *Google Play*. Data yang diperoleh merupakan data dari database website *Google Play*, yaitu data yang berupa ulasan pengguna *Zoom Cloud Meetings* sebanyak 96.122 ulasan.

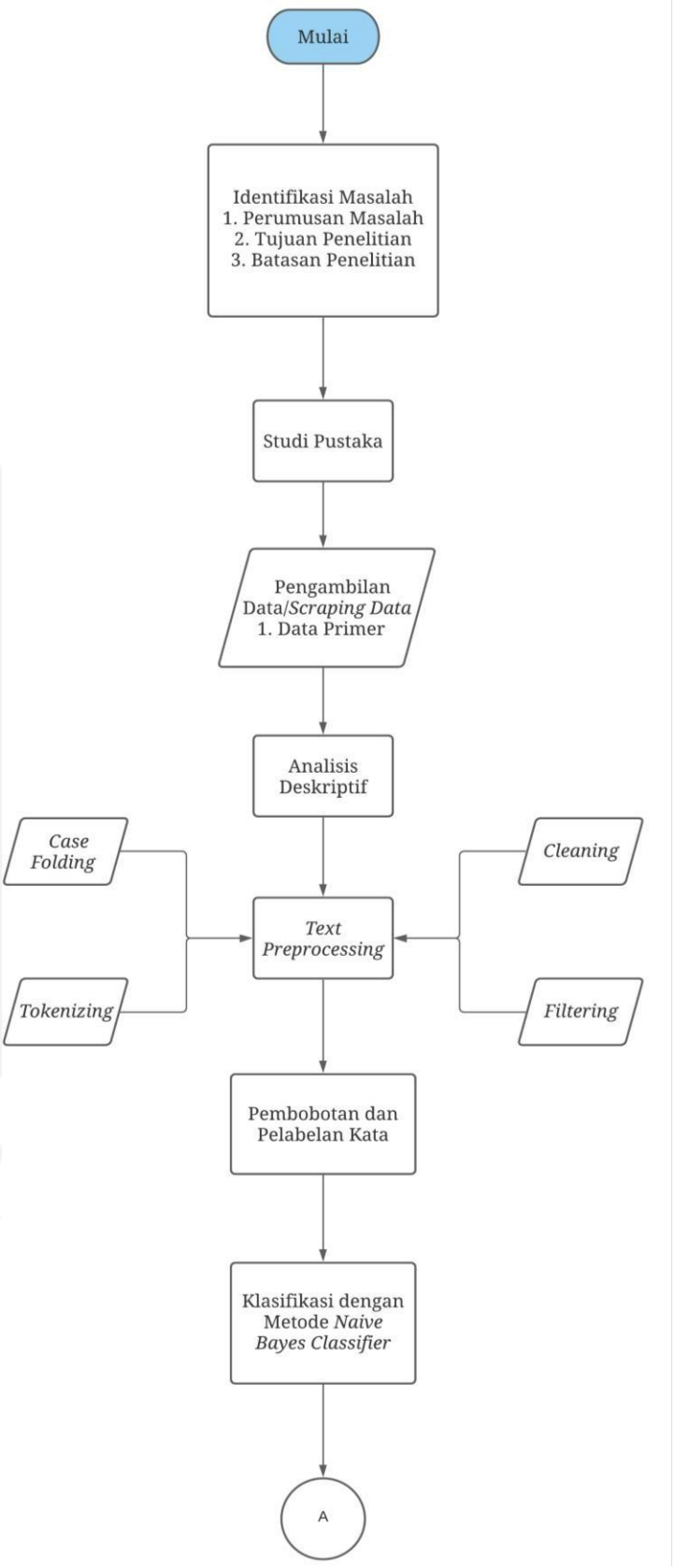
3.5 Metode Analisis Data

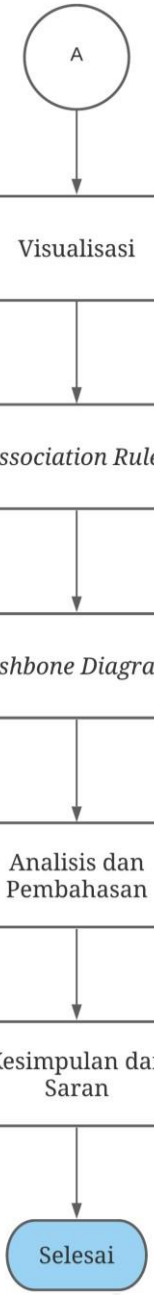
Dalam penelitian ini, peneliti menggunakan *software appfollow.id, Python* 3.8.6 dan *Microsoft Excel* 2010. Adapun metode analisis data yang digunakan dalam penelitian ini yaitu:

1. Analisis Deskriptif, digunakan untuk memberikan gambaran umum ulasan pengguna *Zoom Cloud Meetings* yang ada pada *Google Play Store*.
2. Metode *text mining* yaitu *text preprocessing* dan *word cloud* yang digunakan untuk menemukan pola kata yang sering muncul didalam suatu ulasan khususnya ulasan negatif.
3. Metode klasifikasi berbasis *machine learning* yaitu *Naïve Bayes Classifier* yang digunakan untuk melakukan klasifikasi ulasan pengguna yang dibagi menjadi kelas positif dan kelas negatif.
4. *Association Rule*, digunakan untuk mengidentifikasi dan membentuk pola kata yang berasosisasi dengan kata lainnya guna mendapatkan informasi yang dianggap penting dan berguna.
5. Diagram *Fishbone*, digunakan untuk mengidentifikasi faktor-faktor penyebab permasalahan yang didapatkan dari ulasan negatif sehingga dapat dilakukan pencarian solusi yang mungkin terhadap masalah yang ada.

3.6 Alur Penelitian

Alur penelitian membahas sistematis mengenai tahapan-tahapan sistematis yang dilakukan dalam penelitian. Adapun alur penelitian yaitu sebagai berikut:


Mulai
Identifikasi Masalah
1. Perumusan Masalah
2. Tujuan Penelitian
3. Batasan Penelitian
Studi Pustaka
Pengambilan Data/*Scraping Data*
1. Data Primer
Analisis Deskriptif
Case Folding
Tokenizing
Text Preprocessing
Cleaning
Filtering
Pembobotan dan Pelabelan Kata
Klasifikasi dengan Metode *Naive Bayes Classifier*
A



Gambar 3.2 *Flowchart* penelitian

Berikut penjelasan dari setiap tahap dalam penelitian yang dilakukan berdasarkan diagram alir penelitian pada Gambar 3.2:

1. Identifikasi Masalah
   a) Perumusan Masalah

      Melakukan identifikasi masalah yang terdapat pada aplikasi *Zoom Cloud Meetings* melalui via *website Google Play*. Mengetahui gambaran umum objek

penelitian dengan melihat *rating* aplikasi dan tanggapan pengguna selama menggunakan aplikasi tersebut baik itu tanggapan positif maupun tanggapan negatif.

b) Tujuan Penelitian

Ide tujuan penelitian didapatkan berdasarkan rumusan masalah pada penelitian dan tujuan penelitian merupakan hasil yang akan dicapai pada akhir penelitian.

c) Batasan Penelitian

Pada penelitian ini terdapat batasan masalah yang bertujuan untuk membatasi ruang lingkup masalah yang terlalu luas sehingga penelitian ini dapat tetap fokus pada hal apa yang ingin diteliti.

2. Studi Pustaka

Studi pustaka dilakukan untuk mencari materi dan teori yang berhubungan dengan penelitian ini dan memudahkan dalam menentukan proses yang akan dilakukan selama penelitian. Studi pustaka meliputi *text mining,* analisis sentimen, klasifikasi *association rule* dan *fishbone diagram.*

3. Pengambilan Data

Pengambilan data dilakukan dengan cara *web scraping data* karena lebih mudah dan lebih efisien dibandingkan dengan cara manual. Pada proses *scraping* menggunakan aplikasi *web scraping* berbasis web yaitu, *appfollow*. Proses pengambilan data dilakukan secara *online* dan penggalian informasi terhadap ulasan-ulasan terhadap penggunaan aplikasi *zoom cloud meetings* Hasil dari pengambilan data dengan *web scraping data* merupakan data primer.

4. Analisis Deskriptif

Pada penelitian ini dilakukan analisis deskriptif untuk mengetahui gambaran umum mengenai pelayanan yang diberikan ataupun performansi aplikasi *zoom cloud meetings* berdasarkan data ulasan pengguna baik ulasan positif maupun ulasan negatif yang telah didapatkan dari hasil *scraping* di *Google Play*.

5. *Text Preprocessing*

Dilakukannya *text preprocessing* bertujuan untuk membersihkan data dan mengubah data menjadi lebih terstruktur. Pada penelitian ini tahapan *text preprocessing* menggunakan *tool python* 3.8.6 dan terdapat empat tahapan yang dilakukan pada *text preprocessing,* yaitu:

a) *Case Folding*

Mengubah semua huruf pada data menjadi huruf *lowercase.*

b) *Cleaning*

Menghilangkan hal yang tidak penting pada data seperti tanda baca, angka,emoji,dll.

c) *Tokenizing*

Tokenizing merupakan proses pemecahan teks kalimat menjadi potongan kata-kata guna membuat kalimat yang terbentuk menjadi lebih bermakna.

d) *Filtering*

Tahap *filtering* adalah tahap mengambil kata-kata penting dari hasil proses *tokenizing*. Penelitian ini menggunakan *stopword* berbahasa Inggris berdasarkan modul NLTK.

6. Pelabelan dan Pembobotan Kata

Pelabelan kelas sentimen dilakukan dengan bantuan *library VADER*. Pelabelan dilakukan dengan membagi data menjadi tiga kelas sentimen yaitu sentimen positif, sentimen netral, dan sentimen negatif.

7. Klasifikasi dengan Metode *Naïve Bayes*

Pada penelitian ini metode *Machine learning* yang digunakan yaitu *Naïve bayes Classifier* yang digunakan untuk melakukan klasifikasi ulasan pengguna yang berbentuk kelas positif maupun negatif. Adapun pada tahap ini menggunakan *python* 3.8.6 untuk melakukan klasifikasi.

8. Visualiasi

Visualisasi merupakan tampilan suatu informasi yang berisi kata yang sering muncul dalam ulasan pengguna baik itu ulasan positif ataupun negatif. Visualisasi ditampilkan dalam bentuk *bar plot* dan *word cloud*

9. *Association Rules*

Association rules dilakukan untuk mencari asosiasi antar kata yang paling sering muncul secara bersamaan agar dapat memperkuat pencarian informasi dan dapat memprediksi suatu penyebab permasalahan dari suatu topik pembicaraan. Asosiasi antar kata dilakukan pada ulasan positif maupun ulasan negatif dengan menggunakan *python* 3.8.6 dan *Ms. Excel* 2010.

10. *Fishbone Diagram*

Akar penyebab permasalahan yang telah didapatkan berdasarkan analisis ulasan negatif dipresentasikan pada diagram *fishbone* dan mengidentifikasi hal apa saja yang perlu dilakukan untuk menangani permasalahan tersebut.

11. Analisis dan Pembahasan

    Pada tahapan ini akan menjelaskan terkait pembahasan gambaran umum ulasan pengguna *Zoom Cloud Meetings*, hasil klasifikasi, hasil *association rules* dan rencana perbaikan berdasarkan akar permasalahan pada *fishbone diagram*.

12. Kesimpulan dan Saran

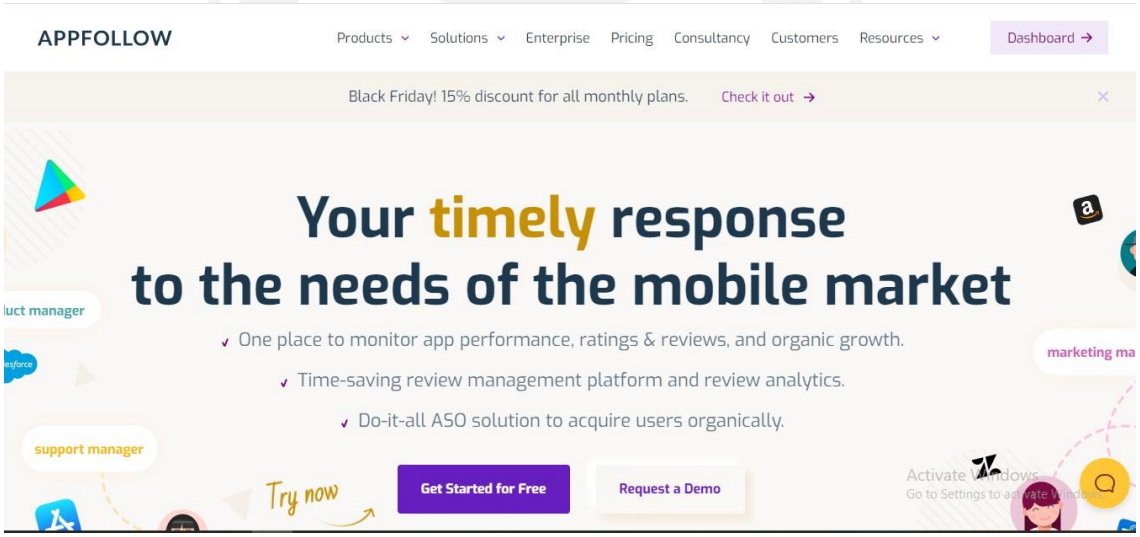
    Tahap terakhir yaitu penarikan kesimpulan dan pemberian saran yang berdasarkan hasil penelitian.

BAB IV

PENGUMPULAN DAN PENGOLAHAN DATA

4.1 Pengumpulan Data

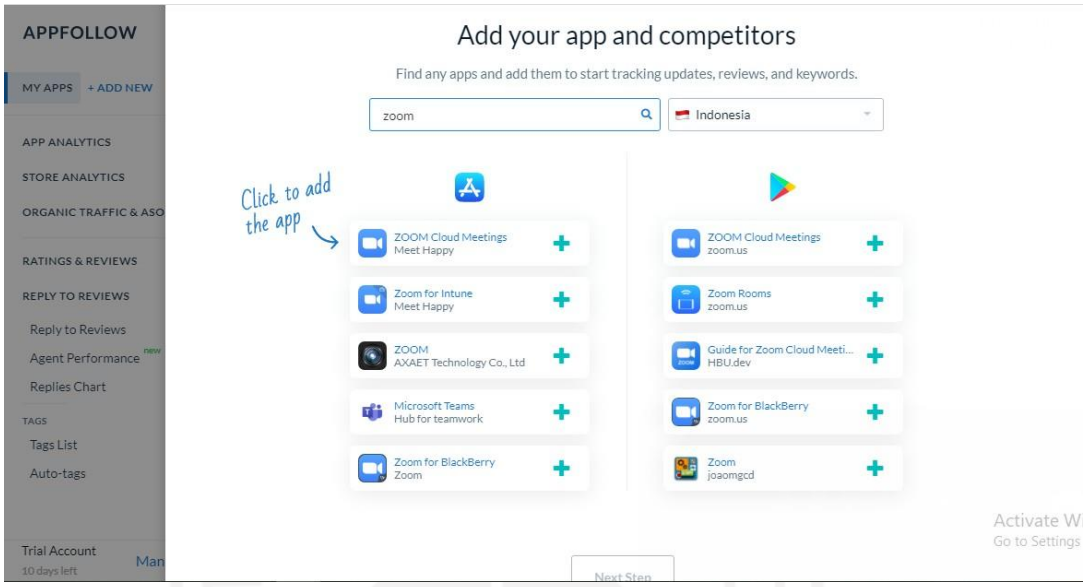
Pada penelitian ini dilakukan tahap pengambilan data berupa ulasan pengguna *Zoom* melalui aplikasinya yaitu *Zoom Cloud Meetings* pada situs *Google Play*. Pengambilan data dilakukan dengan teknik *web scraping* menggunakan aplikasi berbasis web. Terdapat berbagai *website* atau aplikasi untuk melakukan *scraping* data. Salah satunya ialah *appfollow*, *appfollow* merupakan sebuah aplikasi *monitoring*, optimisasi toko, dan *customer support platform*. *Appfollow* dapat digunakan untuk mengestraksi data secara gratis (dalam 10 hari). Data yang diekstraksi dapat diekspor menjadi *file* dengan format .xlsx (*Ms. Excel)* maupun .CSV.



Gambar 4.1 *Interface* Beranda *Appfollow*

Sumber: https://appfollow.io/ (2020)

Tahap pertama yang harus dilakukan adalah mengakses ke situs *appfollow*, yaitu https://appfollow.id dan *sign up* atau membuat akun terlebih dahulu. Pada tahap proses pembuatan akun, *user* baru diharuskan memilih aplikasi apa saja untuk dilakukan *scraping data*.



Gambar 4.2 Pemilihan Aplikasi

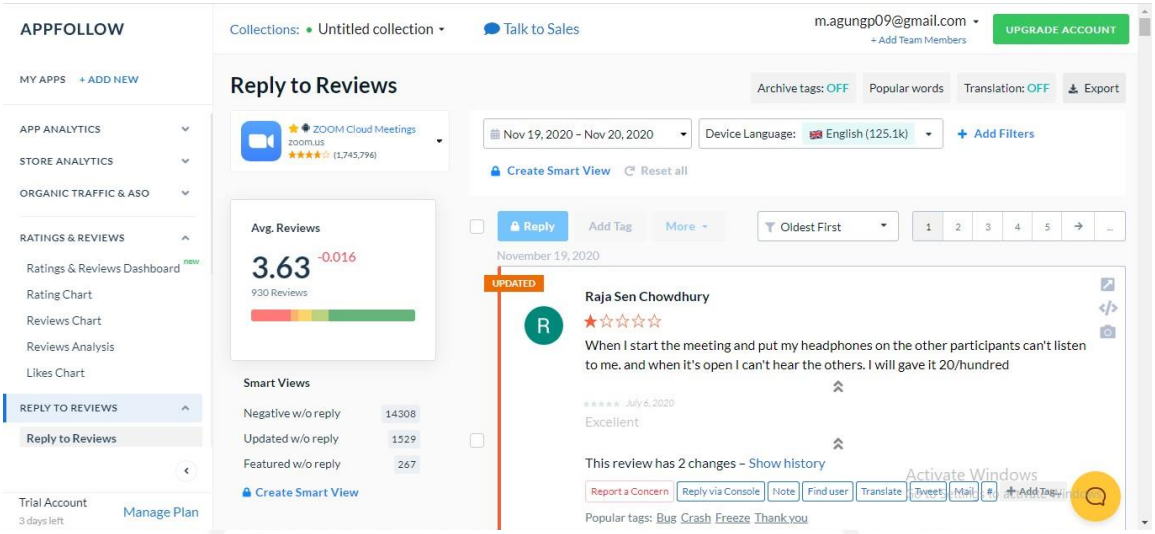
Sumber: https://appfollow.io/ (2020)

Setelah melakukan pemilihan aplikasi maka akan masuk ke tampilan seperti pada gambar 4.3, yaitu *dashboard rating* & ulasan. Selanjutnya klik *Reply to Reviews* pada bagan kiri dan akan diarahkan ke tampilan ulasan-ulasan pengguna *zoom cloud meetings* seperti yang ditunjukkan pada gambar 4.4.



Gambar 4.3 *Dashboard Rating* & Ulasan

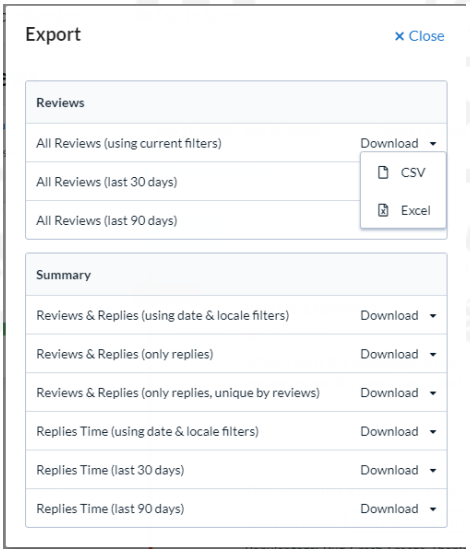
Sumber: https://appfollow.io/ (2020)



Gambar 4.4 *Interface Reply to Reviews*

Sumber: https://appfollow.io/ (2020)

Langkah selanjutnya adalah memilih rentang waktu pengambilan data dan memilih bahasa apa yang akan digunakan untuk di*scraping*. Selanjutnya dapat dilakukan *scraping* dengan mengklik *export* pada kanan atas. Sebelum datanya diekstraksi, pada gambar 4.5 menampilkan data yang akan dieksport dapat diekstraksi menjadi *file* xlsx atau CSV dan dapat ditentukan jangka waktu ulasannya sesuai yang diinginkan.

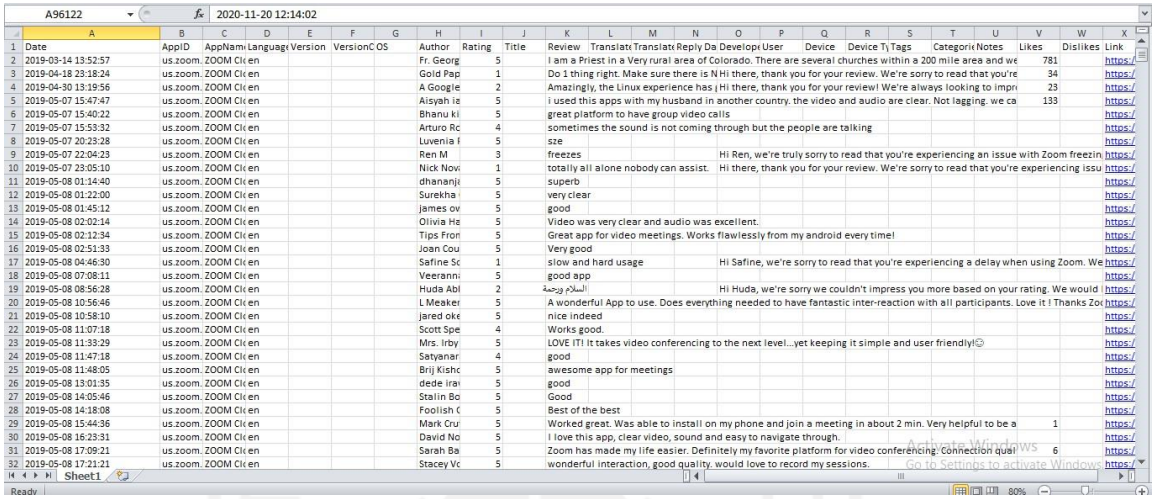


Gambar 4.5 *Export Data*

Sumber: https://appfollow.io/ (2020)

Adapun pada penelitian ini hasil *scraping* data disimpan dalam bentuk *file* .xlsx. Pada gambar 4.6 menunjukkan hasil *scraping* yang telah diekstraksi dan dapat dilihat bahwa

terdapat beberapa kolom yang kurang penting dan dapat diabaikan. Maka dari itu, pada penelitian ini hanya dua aspek/kolom saja yang akan digunakan untuk tahap pengolahan data yaitu, terdiri dari komentar dan *rating*.



Date	AppID	AppNam	Languag	Version	VersionC	OS	Author	Rating	Title	Review	Translat	Translat	Reply Da	Develop	User	Device	Device Ty	Tags	Categori	Notes	Likes	Dislikes	Link
2019-03-14 13:52:57	us.zoom.	ZOOM Cl	en				Fr. Georg	5		I am a Priest in a Very rural area of Colorado. There are several churches within a 200 mile area and we											781		https:/
2019-04-18 23:18:24	us.zoom.	ZOOM Cl	en				Gold Pap	1		Do 1 thing right. Make sure there is N	Hi there, thank you for your review. We're sorry to read that you're										34		https:/
2019-04-30 13:19:56	us.zoom.	ZOOM Cl	en				A Google	2		Amazingly, the Linux experience has j	Hi there, thank you for your review! We're always looking to impr										23		https:/
2019-05-07 15:47:47	us.zoom.	ZOOM Cl	en				Aisyah ia	5		i used this apps with my husband in another country. the video and audio are clear. Not lagging. we ca											133		https:/
2019-05-07 15:40:22	us.zoom.	ZOOM Cl	en				Bhanu ki	5		great platform to have group video calls													https:/
2019-05-07 15:53:32	us.zoom.	ZOOM Cl	en				Arturo Rc	4		sometimes the sound is not coming through but the people are talking													https:/
2019-05-07 20:23:28	us.zoom.	ZOOM Cl	en				Luvenia F	5		sze													https:/
2019-05-07 22:04:23	us.zoom.	ZOOM Cl	en				Ren M	3		freezes				Hi Ren, we're truly sorry to read that you're experiencing an issue with Zoom freezin									https:/
2019-05-07 23:05:10	us.zoom.	ZOOM Cl	en				Nick Nov	1		totally all alone nobody can assist.	Hi there, thank you for your review. We're sorry to read that you're experiencing issu												https:/
2019-05-08 01:14:40	us.zoom.	ZOOM Cl	en				dhananja	5		superb													https:/
2019-05-08 01:22:00	us.zoom.	ZOOM Cl	en				Surekha	5		very clear													https:/
2019-05-08 01:45:12	us.zoom.	ZOOM Cl	en				james ov	5		good													https:/
2019-05-08 02:02:14	us.zoom.	ZOOM Cl	en				Olivia Ha	5		Video was very clear and audio was excellent.													https:/
2019-05-08 02:12:34	us.zoom.	ZOOM Cl	en				Tips Fron	5		Great app for video meetings. Works flawlessly from my android every time!													https:/
2019-05-08 02:51:33	us.zoom.	ZOOM Cl	en				Joan Cou	5		Very good													https:/
2019-05-08 04:46:30	us.zoom.	ZOOM Cl	en				Safine Sc	1		slow and hard usage				Hi Safine, we're sorry to read that you're experiencing a delay when using Zoom. We									https:/
2019-05-08 07:08:11	us.zoom.	ZOOM Cl	en				Veeranni	5		good app													https:/
2019-05-08 08:56:28	us.zoom.	ZOOM Cl	en				Huda Abl	2		السلام ورحمة				Hi Huda, we're sorry we couldn't impress you more based on your rating. We would									https:/
2019-05-08 10:56:46	us.zoom.	ZOOM Cl	en				L Meaker	5		A wonderful App to use. Does everything needed to have fantastic inter-reaction with all participants. Love it ! Thanks Zo													https:/
2019-05-08 10:58:10	us.zoom.	ZOOM Cl	en				jared oke	5		nice indeed													https:/
2019-05-08 11:07:18	us.zoom.	ZOOM Cl	en				Scott Spe	4		Works good.													https:/
2019-05-08 11:33:29	us.zoom.	ZOOM Cl	en				Mrs. Irby	5		LOVE IT! It takes video conferencing to the next level...yet keeping it simple and user friendly!😊													https:/
2019-05-08 11:47:18	us.zoom.	ZOOM Cl	en				Satyanar	4		good													https:/
2019-05-08 11:48:05	us.zoom.	ZOOM Cl	en				Brij Kishc	5		awesome app for meetings													https:/
2019-05-08 13:01:35	us.zoom.	ZOOM Cl	en				dede ira	5		good													https:/
2019-05-08 14:05:46	us.zoom.	ZOOM Cl	en				Stalin Bo	5		Good													https:/
2019-05-08 14:18:08	us.zoom.	ZOOM Cl	en				Foolish (	5		Best of the best													https:/
2019-05-08 15:44:36	us.zoom.	ZOOM Cl	en				Mark Cru	5		Worked great. Was able to install on my phone and join a meeting in about 2 min. Very helpful to be a											1		https:/
2019-05-08 16:23:31	us.zoom.	ZOOM Cl	en				David No	5		I love this app, clear video, sound and easy to navigate through.													https:/
2019-05-08 17:09:21	us.zoom.	ZOOM Cl	en				Sarah Ba	5		Zoom has made my life easier. Definitely my favorite platform for video conferencing. Connection quali											6		https:/
2019-05-08 17:21:21	us.zoom.	ZOOM Cl	en				Stacey Vc	5		wonderful interaction, good quality. would love to record my sessions.													https:/

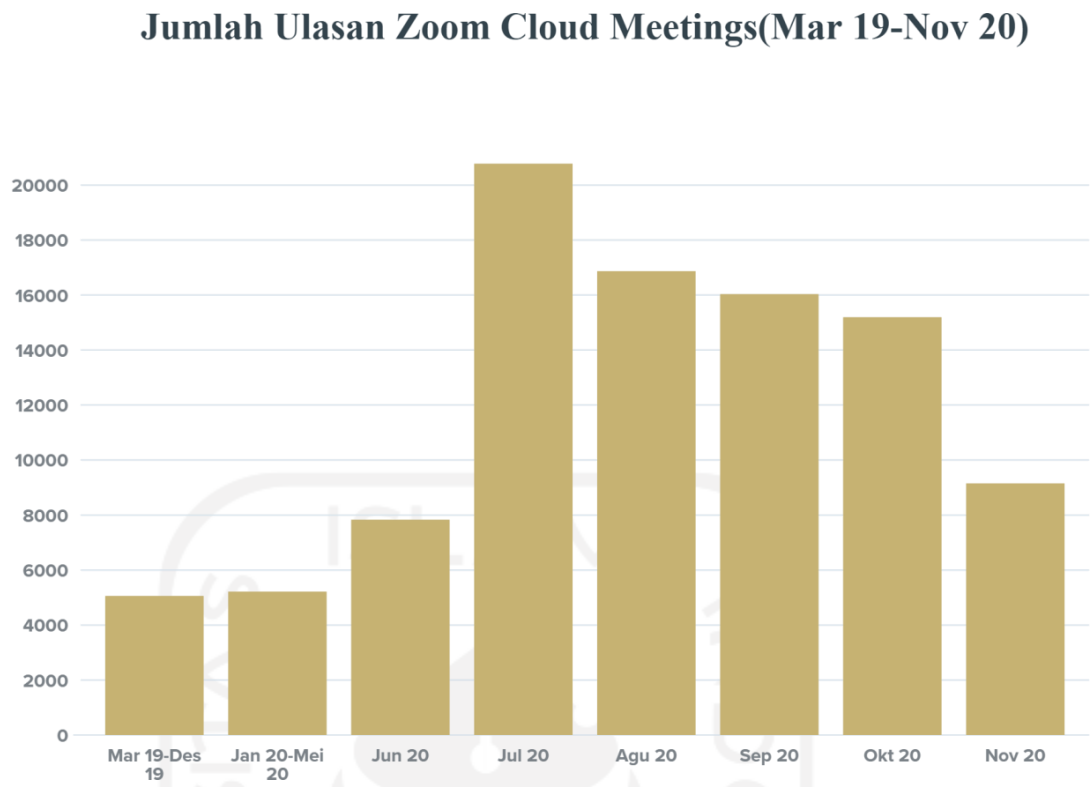
Gambar 4.6 Hasil *Scraping* dari *AppFollow*

4.2 Pengolahan Data

Data yang telah diekstraksi oleh *appfollow* selanjutnya dilakukan tahap pengolahan agar dapat menemukan informasi penting dari data tersebut. Pengolahan data dilakukan pada ulasan pengguna *Zoom* dalam Bahasa Inggris dengan sampel data sebanyak 96.122 ulasan.

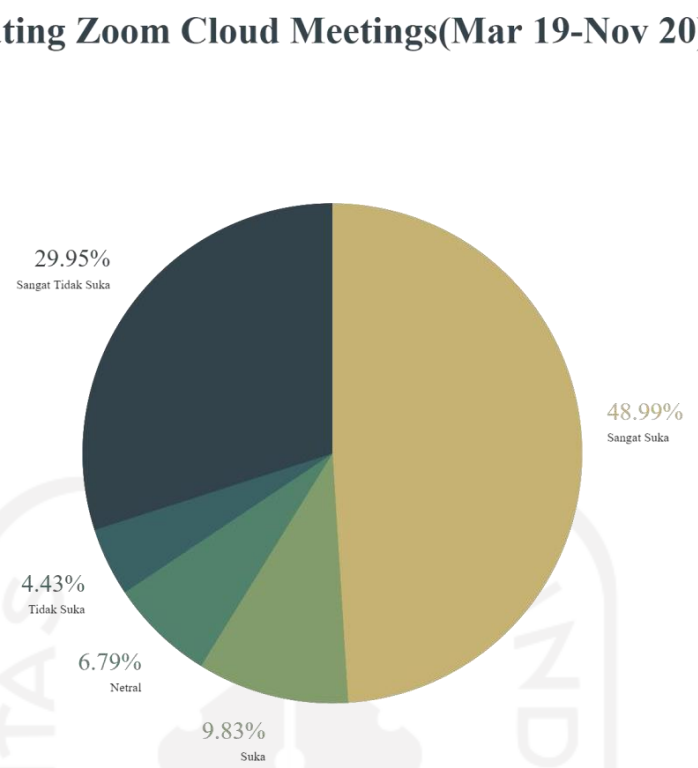
4.2.1 Analisis Deskriptif

Pada penelitian ini dilakukan analisis deskriptif untuk mengetahui gambaran umum mengenai pelayanan yang diberikan ataupun performansi aplikasi *zoom cloud meetings* berdasarkan data ulasan pengguna baik ulasan positif maupun ulasan negatif yang telah didapatkan dari hasil *scraping* di *Google Play*. Ada dua aspek yang ditampilkan pada analisis deskriptif ini yaitu, jumlah ulasan berdasarkan urutan waktu (Maret 2019 - November 2020) dan persentase setiap *rating* yang diberikan oleh pengguna. Berikut ini merupakan *rating* yang menampilkan penilaian pengguna terhadap aplikasi *zoom cloud meetings* di *google play*.



Gambar 4.7 Jumlah Ulasan Pengguna

Berdasarkan Gambar 4.7, dengan rentang waktu pengambilan data selama 14 Maret 2019 sampai dengan 20 November 2020 diperoleh hasil grafik yang awalnya naik kemudian turun. Adapun pada bulan Maret-Desember 2019 hanya terdapat 5057. Pada bulan Januari-Mei 2020 terdapat 5216 ulasan, pada bulan Juni ulasan yang masuk meningkat menjadi 7830 ulasan. Selanjutnya pada bulan Juli ulasan meningkat drastis menjadi 20.775, hal ini diperkirakan terjadi karena pada bulan-bulan pertengahan tahun 2020 semua negara sudah menerapkan sistem *Work From Home* (WFH) dan karantina di rumah yang dimana masyarakat hanya dapat melakukan aktivitas secara daring baik itu bekerja, belajar, silaturahmi *online*, dan lain sebagainya, sehingga penggunaan jasa layanan aplikasi berbasis *video conference* termasuk didalamnya *zoom cloud meetings* juga turut meningkat.



Gambar 4 8 Persentase *Rating*

Berdasarkan Gambar 4.8 diketahui bahwa terdapat 5 kategori *rating* yang berbeda berdasarkan penilaian pengguna yaitu, bintang dengan nilai 1 sampai dengan 5. Bintang 1 termasuk dalam kategori "Sangat Tidak Suka", bintang 2 "Tidak Suka", bintang 3 "Netral", bintang 4 "Suka" dan bintang 5 "Sangat Suka". Dari Gambar 4.7 diatas dapat diketahui bahwa hampir setengah pengguna aplikasi *Zoom Cloud Meetings* mempunyai penilaian yang baik terhadap layanan tersebut. Hal ini terbukti berdasarkan jumlah penilaian pengguna dari 96.122 ulasan, sebanyak 48.99% atau sebanyak 47.086 pengguna memberikan *rating* 5 atau sangat suka. Selanjutnya untuk persentase terbanyak kedua adalah 29.95% atau 28.791 pengguna yang mana memberikan *rating* 1 atau sangat tidak suka. Kemudian diikuti oleh *rating* 4 sebanyak 9.83% dengan 9.451 pengguna, *rating* 3 sebanyak 6.79% dengan 6.531 pengguna, dan *rating* 2 sebanyak 4.43% dengan 4.263 pengguna.

4.2.2 *Text Preprocessing*

Pada tahap *preprocessing*, dilakukan pembersihan data menggunakan metode *text mining*. Hal ini disebabkan data ulasan hasil *scraping* memiliki format yang tidak

terstruktur, sehingga informasi yang ada di dalam data tidak bisa diekstrak secara langsung. Maka dari itu, diperlukan tahapan *text preprocessing* yang merupakan suatu tahapan yang bertujuan untuk menyeragamkan bentuk kata, menghilangkan karakter-karakter selain huruf, dan mengurangi volume kosakata sehingga data akan lebih terstruktur. Adapun *preprocessing* data pada penelitian ini dilakukan dengan menggunakan *tools* yaitu *Python* 3.8.6 dan *Ms. Excel*. Tahapan *preprocessing* memiliki beberapa tahapan yang akan dilakukan, antara lain sebagai berikut :

1. *Cleaning*

   Dilakukannya tahapan *cleaning* ini bertujuan untuk membersihkan data ulasan dari hal yang tidak penting seperti tanda baca, angka, dan sebagainya. Dalam melakukan proses *cleaning* tersebut dilakukan 3 tahapan, di antaranya ialah:

   - Menghapus angka

   Tabel 4.1 Contoh Penghapusan Angka Pada Data

Input	***Output***
Its great until 2 people talk at the same time. Then the world explodes. The next 10 seconds consists of digital noise and sentence fragments, followed by a scream usually, then an obscenity.	Its great until people talk at the same time. Then the world explodes. The next seconds consists of digital noise and sentence fragments, followed by a scream usually, then an obscenity.

   - Menghapus tanda baca

   Tabel 4.2 Contoh Penghapusan Tanda Baca Pada Data

Input	***Output***
Always dependable. Zoom has made so much possible during this time of shut downs & restrictions.	Always dependable Zoom has made so much possible during this time of shut downs restrictions

   - Menghapus *emoji*

Tabel 4.3 Contoh Penghapusan *Emoji* Pada Data

Input	***Output***
very Nice App 🔥 for the different types of meeting GGGGGGGG	very Nice App for the different types of meeting

Case folding merupakan tahapan penyeragaman bentuk huruf dimana dalam proses ini hanya menerima huruf latin antara "a" sampai "z". Karakter lain selain huruf dianggap sebagai *delimiter* (pembatas) sehingga karakter tersebut akan dihapus dari dokumen/data. Penyeragaman dilakukan dengan cara mengubah isi data teks menjadi huruf kecil secara keseluruhan (dari "a" sampai dengan "z"). Hal ini bertujuan agar kata yang ditulis dengan huruf awal kapital dan huruf non kapital tidak terdeteksi memiliki arti yang berbeda dan tidak terhitung sebagai redundan. Berikut merupakan contoh pengaplikasian *case folding* yang dapat dilihat pada Tabel 4.4 berikut :

Tabel 4.4 Contoh *Case Folding* Pada Data

Input	***Output***
Nice work for you to the same Day installation of the year and I will be in the future of our games are b	nice work for you to the same day installation of the year and i will be in the future of our games are b

2. *Tokenization*

Tokenization merupakan proses pemecahan teks kalimat menjadi potongan kata-kata guna membuat kalimat yang terbentuk menjadi lebih bermakna yang disebut token. *Tokenization* bertujuan untuk mendapatkan potongan kata berupa entitas yang memiliki nilai dalam penyusunan matriks dokumen pada proses selanjutnya. *Tokenization* dapat memudahkan proses perhitungan kemunculan kata dalam data. Berikut merupakan contoh penerapan *tokenization* pada kalimat yang tertera pada Tabel 4.5 :

Tabel 4.5 Contoh Penerapan *Tokenization* Pada Data

Input	***Output***
in middle of the meeting it screen disappeared and sound quality is also bad sometimes	“in” “middle” “of” “the” “meeting” “it” “screen” “disappeared” “and” “sound” “quality” “is” “also” “bad” “sometimes”

3. *Filtering*

Tahap *filtering* adalah tahap mengambil kata-kata penting dari hasil proses *tokenization*. Proses *filtering* dapat menggunakan algoritma *stoplist* (membuang kata yang kurang penting) atau *wordlist* (menyimpan kata penting). Penelitian ini menggunakan *stopword* berbahasa Inggris berdasarkan modul NLTK. *Stopword / stoplist* adalah kata-kata yang tidak memiliki makna penting yang dapat dibuang dalam pendekatan *bag-of-words*. Contoh *stopword* dalam Bahasa Inggris adalah “*the*”, “*and*”, “*to*”, “*i*”, “*you*”, ”*was*” dan lain-lain. Selain itu, proses *filtering* juga dapat dilakukan secara manual, yakni dengan menghapus kata-kata yang tidak terdapat dalam daftar *stopword* dan kata-kata yang kurang berpengaruh secara signifikan terhadap hasil analisis. Adapun terdapat beberapa kata yang ditambahkan secara manual dikarenakan tidak terdapat didalam *stopword.* Contoh dari *stopwords* seperti ”*app*”, “*ever*”, “*bad*”, “*worst*”, dan lain sebagainya. Berikut merupakan contoh *filtering* pada kalimat ulasan seperti pada Tabel 4.6:

Tabel 4.6 Contoh Penerapan *Filtering* Data

Input	***Output***
“in” “middle” “of” “the” “meeting” “it” “screen” “disappeared” “and” “sound” “quality” “is” “also” “bad” “sometimes”	“middle” “meeting” “screen” “disappeared” “sound” “quality” “also” “bad” “sometimes”

4.2.3 Representasi Model

Representasi model dilakukan dengan membuat model dari data yang masih berbentuk kata-kata agar dapat diolah dan dihitung. Representasi dilakukan dengan melakukan perhitungan skor kata berdasarkan kamus *lexicon* dan kemudian memberi label kelas sentimen berdasarkan skor yang ada. Skor sentimen dihitung berdasarkan jumlah kata yang terdeteksi pada kamus lexicon.

a) Pelabelan Kelas Sentimen

Tahap berikutnya setelah mendapatkan skor terhadap sentimen adalah melakukan pelabelan kelas sentimen. Pelabelan kelas sentimen dilakukan dengan bantuan *library VADER. VADER* merupakan salah satu *library* yang populer pada analisis sentimen (Vaish, 2018). Pelabelan dilakukan dengan membagi data menjadi tiga kelas sentimen yaitu sentimen positif, sentimen netral, dan sentimen negatif dengan ketentuan skor sebagai berikut (Hutto & Gilbert, 2014) :

- Sentimen Positif = >= 0,05
- Sentimen Netral = > -0,05 dan < 0,05
- Sentimen Negatif = <= -0,05

Kelas sentimen positif merupakan pernyataan yang mengandung ungkapan pujian, rasa terimakasih, dan lain sebagainya. Adapun untuk kelas sentimen netral tidak mengandung kata yang bersifat positif maupun negatif, atau yang bersifat seimbang. Kelas sentimen negatif merupakan pernyataan yang bersifat ketidakpuasan pengguna, penghinaan dan segala sesuatunya yang bersifat negatif. Adapun hasil pelabelan kelas sentimen diperoleh jumlah data seperti berikut :

Tabel 4.7 Hasil Pelabelan Kelas Sentimen berbasis Kamus *Lexicon*

Ulasan	Skor Sentimen	Kelas
great platform to have group video calls	neg: 0.000 neu: 0.594 pos: 0.406 compound: 0.6249	Positif
This is very bad how can you just show off symbol	neg: 0.378 neu: 0.622	Negatif

Ulasan	Skor Sentimen	Kelas
in participants worst app	pos: 0.000 compound: -0.8357	

Pada *library* VADER setiap kata dalam Bahasa Inggris memiliki bobotnya tersendiri. Penentuan kelas sentimen berdasarkan hasil *compound* yang telah didapatkan. *Compound* pada VADER merupakan sebuah metrik yang menghitung jumlah semua *rating lexicon* yang telah dinormalisasi antara -1 (paling negatif paling ekstrim) dan +1 (paling positif paling ekstrim) (Hutto & Gilbert, 2014).

Tabel 4.8 Perbandingan Jumlah Data pada Kelas Sentimen

Kelas Sentimen	Jumlah Ulasan
Positif	55.413
Netral	18.784
Negatif	20.041

Berdasarkan Tabel 4.8, hasil pelabelan kelas sentimen menunjukkan bahwa jumlah ulasan positif memiliki frekuensi yang lebih tinggi dibandingkan dengan jumlah ulasan netral maupun negatif. Jumlah ulasan positif sebanyak 55.413 ulasan, ulasan netral sebanyak 18.784 dan ulasan negatif sebanyak 20.041 ulasan. Klasifikasi yang akan digunakan pada penelitian ini hanya data dengan sentimen positif dan negatif. Sedangkan untuk sentimen netral tidak dimasukkan ke dalam tahap klasifikasi dikarenakan pada penelitian ini berfokus mencari permasalahan yang umumnya didapatkan dari ulasan negatif pengguna.

4.2.4 Pembagian Data Latih dan Data Uji

Didalam suatu *dataset* atau himpunan data terdapat dua jenis data, yaitu *data training* dan *data testing*. *Data training* oleh algoritma klasifikasi digunakan untuk membangun sebuah model data *testing* dan juga *data training* digunakan untuk perhitungan probabilitas dari data berdasarkan data pembelajaran yang dilakukan (Wijayatun &

Nugroho, 2016). *Data training* merupakan data yang sebelumnya sudah ada sesuai dengan fakta. Sedangkan *data testing* merupakan data yang akan atau sedang terjadi dan dipergunakan sebagai bahan uji yang sebelumnya sudah didapatkan pada *data training* (Wijayatun & Nugroho, 2016). Data ini digunakan untuk mengukur sejauh mana klasifikasi berhasil melakukan klasifikasi dengan benar. Data yang akan digunakan untuk *data training* dan *data testing* adalah data yang telah diberikan label kelas.

Adapun pada penelitian ini dilakukan empat kali percobaan dengan perbandingan antara data latih dan data uji yang berbeda, antara lain sebagai berikut :

1) Perbandingan data latih sebesar 60% dan data uji 40% :

Tabel 4.9 Pembagian Data Latih dan Data Uji 1

Klasifikasi	Jumlah	Data Latih	Data Uji
Positif	55.413	33.248	22.165
Negatif	20.041	12.025	8.016
Total	75.454	45.273	30.181

Berdasarkan perbandingan 60% : 40%, dari total 75.454 ulasan berbahasa Inggris digunakan 45.273 data latih dan 30.181 data uji.

2) Perbandingan data latih sebesar 70% dan data uji 30% :

Tabel 4.10 Pembagian Data Latih dan Data Uji 2

Klasifikasi	Jumlah	Data Latih	Data Uji
Positif	55.413	38.789	16.624
Negatif	20.041	14.029	6.012
Total	75.454	52.818	22.636

Berdasarkan perbandingan 70% : 30%, dari total 75.454 ulasan berbahasa Inggris digunakan 52.818 data latih dan 22.636 data uji.

3) Perbandingan data latih sebesar 80% dan data uji 20% :

Tabel 4.11 Pembagian Data Latih dan Data Uji 3

Klasifikasi	Jumlah	Data Latih	Data Uji
Positif	55.413	44.330	11.083
Negatif	20.041	16.033	4.008
Total	75.454	60.363	15.091

Berdasarkan perbandingan 80% : 20%, dari total 75.454 ulasan berbahasa Inggris digunakan 60.363 data latih dan 15.091 data uji.

4) Perbandingan data latih sebesar 85% dan data uji 15% :

Tabel 4.12 Pembagian Data Latih dan Data Uji 4

Klasifikasi	Jumlah	Data Latih	Data Uji
Positif	55.413	47.101	8.312
Negatif	20.041	17.035	3.006
Total	75.454	64.136	11.318

Berdasarkan perbandingan 85% : 15%, dari total 75.454 ulasan berbahasa Inggris digunakan 64.136 data latih dan 11.318 data uji.

4.1.2 Klasifikasi dengan Metode *Naïve Bayes Classifier*

Setelah dilakukan penentuan terkait data latih dan data uji selanjutnya dilakukan klasifikasi dengan metode *Naïve Bayes Classifier*. Analisis *Naïve Bayes Classifier* dilakukan dengan membuat *confusion matrix* agar dapat diketahui nilai akurasi, recall dan presisi. Pada penelitian ini *confusion matrix* dibuat dengan menggunakan aplikasi *python* 3.8.6. Berdasarkan pembagian data latih dan data uji yang berbeda, didapatkan hasil *confusion matrix* dengan nilai akurasi, presisi, dan *recall* yang berbeda pula yang dapat dilihat pada tabel berikut :

1. Data Latih 60% data Uji 40%

Tabel 4.13 Hasil Klasifikasi 1

Prediksi	Aktual		*Class Precision*
	Negatif	Positif	
Negatif	3101	1112	92,35%

Positif	257	5530	83,26%
Class Recall	73,61%	95,56%	
Akurasi		86,31%	

Berdasarkan *confusion matrix* pada Tabel 4.13 dengan menggunakan metode *Naïve Bayes Classifier*, diketahui bahwa prediksi ulasan negatif berjumlah 4213 ulasan. Dari 4213 prediksi ulasan negatif, terdapat kesesuaian bahwa ulasan negatif yang termasuk aktual negatif adalah sebanyak 3101 data yang telah terklasifikasi dengan benar, dan terdapat ketidaksesuaian karena kebenarannya ulasan adalah positif (aktual positif) sebanyak 1112 ulasan. Sehingga pada prediksi negatif didapatkan *Class Precision* sebanyak 92,35%. Pada prediksi positif diketahui terdapat 5787 ulasan dengan jumlah kesesuaian bahwa ulasan yang aktual positif sebanyak 5530 yang telah terklasifikasi dengan benar dan terdapat ketidaksesuaian atau dengan kata lain terdapat kesalahan prediksi sebanyak 257 ulasan positif yang masuk dalam ulasan negatif. Sehingga didapatkan nilai presisi untuk kelas positif sebesar 83,26%. Selanjutnya dari hasil *confusion matrix* tersebut diperoleh tingkat akurasi sebesar 86,31%, yang artinya dari 10.000 data ulasan yang diuji, terdapat 8631 ulasan yang diklasifikasikan secara benar oleh model *Naïve Bayes Classifier*. Sehingga *error rate* (100% - 86,31%) sebesar 13,69%.

2. Data Latih 70% data Uji 30%

Tabel 4.14 Hasil Klasifikasi 2

Prediksi	Aktual		*Class Precision*
	Negatif	Positif	
Negatif	2337	831	91,83%
Positif	208	4124	83,23%
Class Recall	73,77%	95,20%	
Akurasi		86,15%	

Berdasarkan *confusion matrix* pada Tabel 4.14 dengan menggunakan metode *Naïve Bayes Classifier*, diketahui bahwa prediksi ulasan negatif berjumlah 3168 ulasan. Dari 3168 prediksi ulasan negatif, terdapat kesesuaian bahwa ulasan negatif yang termasuk aktual negatif adalah sebanyak 2337 data yang telah terklasifikasi dengan benar, dan

terdapat ketidaksesuaian karena kebenarannya ulasan adalah positif (aktual positif) sebanyak 831 ulasan. Sehingga pada prediksi negatif didapatkan *Class Precision* sebanyak 91,83%. Pada prediksi positif diketahui terdapat 4332 ulasan dengan jumlah kesesuaian bahwa ulasan yang aktual positif sebanyak 4124 yang telah terklasifikasi dengan benar dan terdapat ketidaksesuaian atau dengan kata lain terdapat kesalahan prediksi sebanyak 208 ulasan positif yang masuk dalam ulasan negatif. Sehingga didapatkan nilai presisi untuk kelas positif sebesar 83,23%. Selanjutnya dari hasil *confusion matrix* tersebut diperoleh tingkat akurasi sebesar 86,15%, yang artinya dari 7500 data ulasan yang diuji, terdapat 6461 ulasan yang diklasifikasikan secara benar oleh model *Naïve Bayes Classifier*. Sehingga *error rate* (100% - 86,15%) sebesar 13,85%.

3. Data Latih 80% data Uji 20%

Tabel 4.15 Hasil Klasifikasi 3

Prediksi	Aktual		*Class Precision*
	Negatif	Positif	
Negatif	1548	537	91,45%
Positif	148	2731	83,57%
Class Recall	74,68%	94,86%	
Akurasi		86,30%	

Berdasarkan *confusion matrix* pada Tabel 4.15 dengan menggunakan metode *Naïve Bayes Classifier*, diketahui bahwa prediksi ulasan negatif berjumlah 2085 ulasan. Dari 2085 prediksi ulasan negatif, terdapat kesesuaian bahwa ulasan negatif yang termasuk aktual negatif adalah sebanyak 1548 data yang telah terklasifikasi dengan benar, dan terdapat ketidaksesuaian karena kebenarannya ulasan adalah positif (aktual positif) sebanyak 537 ulasan. Sehingga pada prediksi negatif didapatkan *Class Precision* sebanyak 91,54%. Pada prediksi positif diketahui terdapat 2879 ulasan dengan jumlah kesesuaian bahwa ulasan yang aktual positif sebanyak 2731 yang telah terklasifikasi dengan benar dan terdapat ketidaksesuaian atau dengan kata lain terdapat kesalahan prediksi sebanyak 148 ulasan positif yang masuk dalam ulasan negatif. Sehingga didapatkan nilai presisi untuk kelas positif sebesar 83,57%. Selanjutnya dari hasil *confusion matrix* tersebut diperoleh tingkat akurasi sebesar 86,30%, yang artinya dari 4964 data ulasan yang diuji, terdapat 4284 ulasan yang diklasifikasikan secara benar

oleh model *Naïve Bayes Classifier*. Sehingga *error rate* (100% - 86,30%) sebesar 13,70%.

4. Data Latih 85% data Uji 15%

Tabel 4.16 Hasil Klasifikasi 4

Prediksi	Aktual		*Class Precision*
	Negatif	Positif	
Negatif	1190	390	91,54%
Positif	110	2060	84,08%
Class Recall	75,32%	94,93%	
Akurasi		86,67%	

Berdasarkan *confusion matrix* pada Tabel 4.16 dengan menggunakan metode *Naïve Bayes Classifier*, diketahui bahwa prediksi ulasan negatif berjumlah 1580 ulasan. Dari 1580 prediksi ulasan negatif, terdapat kesesuaian bahwa ulasan negatif yang termasuk aktual negatif adalah sebanyak 1190 data yang telah terklasifikasi dengan benar, dan terdapat ketidaksesuaian karena kebenarannya ulasan adalah positif (aktual positif) sebanyak 390 ulasan. Sehingga pada prediksi negatif didapatkan *Class Precision* sebanyak 91,54%. Pada prediksi positif diketahui terdapat 2170 ulasan dengan jumlah kesesuaian bahwa ulasan yang aktual positif sebanyak 2060 yang telah terklasifikasi dengan benar dan terdapat ketidaksesuaian atau dengan kata lain terdapat kesalahan prediksi sebanyak 110 ulasan positif yang masuk dalam ulasan negatif. Sehingga didapatkan nilai presisi untuk kelas positif sebesar 84,08%. Selanjutnya dari hasil *confusion matrix* tersebut diperoleh tingkat akurasi sebesar 86,67%, yang artinya dari 3750 data ulasan yang diuji, terdapat 3250 ulasan yang diklasifikasikan secara benar oleh model *Naïve Bayes Classifier*. Sehingga *error rate* (100% - 86,67%) sebesar 13,33%.

4.2.5 Visualisasi

Visualisasi dilakukan pada keseluruhan data, baik pada ulasan positif maupun ulasan negatif. Tujuannya dilakukan visualisasi adalah untuk memperoleh informasi secara menyeluruh mengenai aspek apa saja yang sering dibicarakan terkait performansi

aplikasi *zoom cloud meetings* melalui ulasan pengguna yang telah dilakukan proses *text preprocessing*.

a. Bar Plot

Bar plot atau *bar chart* merupakan grafik yang berfungsi menunjukkan frekuensi atau fraksi dari kategori data. *Bar plot* merupakan salah satu metode yang bertujuan untuk memvisualisasikan data teks.

a. Ulasan Positif

Pada visualisasi ulasan positif menggunakan *bar plot* ini digunakan data ulasan dari hasil *text preprocessing* dan data yang sudah dilakukan *sorting* menggunakan *tool ms.excel*. Data ulasan positif diekstraksi berulang kali agar mendapatkan kata yang paling sering digunakan oleh pengguna *zoom*, selanjutnya visualisasi *barplot* diidentifikasi berdasarkan frekuensi kata pada ulasan positif. Berikut merupakan hasil ekstraksi informasi berupa visualisasi dalam bentuk *bar plot* berdasarkan ulasan positif pengguna *zoom*:



Gambar 4.9 *Bar Plot* Ulasan Positif

Berdasarkan Gambar 4.09 diketahui bahwa terdapat beberapa kata yang paling sering muncul, dalam hal ini yang ditampilkan hanya 10 kata dengan frekuensi terbanyak antara lain : kata "*meeting*" sebanyak 5441 kali, kata "*class*" sebanyak

2878 kali, kata "*video*" sebanyak 2667 kali, kata "*online*" sebanyak 2494 kali, kata "*time*" sebanyak 1483 kali kata "*problem*" sebanyak 1442 kali, kata "*work*" sebanyak 1386 kali, kata "*connect*" sebanyak 1140 kali, kata "*audio*" sebanyak 985 kali, dan kata "*update*" sebanyak 945 kali.

b. Ulasan Negatif

Berdasarkan jumlah total ulasan pengguna *zoom cloud meetings* yaitu, sebanyak 96.122, diketahui bahwa sekitar 30% diantaranya termasuk dalam ulasan negatif atau sebanyak 28.791 ulasan. Data ulasan negatif diekstraksi berulang kali agar mendapatkan kata yang paling sering digunakan oleh pengguna *zoom*, selanjutnya visualisasi *barplot* diidentifikasi berdasarkan frekuensi kata pada ulasan negatif. Selain itu juga didasarkan pada relevansi kata dengan topik yang mengacu pada ulasan negatif. Berikut adalah visualisasi hasil ekstraksi informasi dalam bentuk *bar plot* yang didapatkan dari ulasan pengguna:



Gambar 4.10 *Bar Plot* Ulasan Negatif

Berdasarkan Gambar 4.10 diketahui bahwa pada ulasan negatif terdapat beberapa kata yang paling sering muncul, dalam hal ini 10 kata dengan frekuensi terbanyak antara lain sebagai berikut: kata "*problem*" sebanyak 3030 kali kata, "*meeting*" sebanyak 2369 kali, kata "*video*" sebanyak 1663 kali, kata "*audio*" sebanyak 1328 kali, kata "*connect*" sebanyak 1272 kali, kata "*class*" sebanyak

1250 kali, kata "*time*" sebanyak 1066 kali, kata "*update*" sebanyak 1000 kali kata "*quality*" sebanyak 991 kali dan kata "*work*" sebanyak 981 kali.

4.2.6 *Association Rule*

Association rule dilakukan untuk mencari asosiasi antar kata yang paling sering muncul secara bersamaan agar dapat memperkuat pencarian informasi dan dapat memprediksi suatu penyebab permasalahan dari suatu topik pembicaraan. Asosiasi antar kata dilakukan pada ulasan positif maupun ulasan negatif dengan bantuan *tool python* 3.8.6 dan *Ms. Excel* 2010.

a. Ulasan Positif

Meeting		Class		Video		Online		Problem	
cloud	9.307	online	13.565	conferenc	20.257	class	13.565	solve	27.127
join	7.487	attend	11.399	participan	11.765	study	5.708	network	12.226
attend	4.880	school	4.947	audio	11.725	school	5.431	fix	11.402
able	3.737	student	4.400	quality	11.628	meeting	2.638	audio	7.400
even	3.408	join	3.889	call	11.113	problem	2.479	update	7.046
people	3.254	help	3.390	clear	7.143	video	1.735	connect	6.758
problem	3.189	problem	3.340	problem	4.789			work	4.803
fix	3.150	time	2.118	meeting	2.422			video	4.789
connect	3.056	work	1.931	class	1.853			time	4.061
time	2.989	meeting	1.898	online	1.735			class	3.340
online	2.638	video	1.853					meeting	3.189
audio	2.546							online	2.479
applicatio	2.519								
video	2.422								
friend	2.230								
work	2.155								
class	1.898								
better	1.640								
experienc	1.593								

Time		Work		Connect		Audio		Update	
connect	5.056	well	22.537	internet	21.495	quality	13.007	new	26.846
problem	4.061	update	7.153	network	13.549	video	11.725	work	7.153
meeting	2.989	audio	5.792	even	10.341	clear	11.557	problem	7.046
video	2.403	problem	4.803	problem	6.289	problem	7.400	video	4.122
class	2.323	video	2.625	time	5.056	work	5.792	meeting	2.335
		meeting	2.155	meeting	3.056	meeting	2.546		
		class	1.931	video	2.970				

Gambar 4.11 Asosiasi Kata Ulasan Positif

Berdasarkan Gambar 4.11 diperoleh asosiasi kata pada ulasan positif. Proses ekstraksi informasi dengan asosiasi dilakukan secara berulang kali dengan cara menyaring kata-kata yang memiliki hubungan dengan kata lain dan didasarkan pada relevansi kata dengan topik yang dibahas.

Berdasarkan asosiasi yang terkait dengan kata "*meeting*" diperoleh informasi bahwa dari segi kegiatan *meeting* aplikasi *zoom cloud meetings* telah memiliki performa yang baik, kualitas yang bagus baik itu dari *audio, video, screen,*dll dan juga *zoom* dinilai baik ketika kita menjadi *host* ataupun peserta *meeting* sehingga pengguna merekomendasi aplikasi ini untuk melakukan *meeting*.

Asosiasi yang terkait dengan kata "class" menunjukkan informasi bahwa aplikasi zoom cloud meetings ini dinilai baik oleh pelajar karena dapat menghadiri kelas secara daring khususnya di masa pandemi ini. Pengguna memberikan saran agar durasi waktu untuk melakukan video tidak dibatasi bagi pengguna gratis.

Asosiasi yang terkait dengan kata "*video*" menunjukkan informasi bahwa menurut pengguna aplikasi *zoom cloud meetings* ini merupakan aplikasi *video conference* terbaik karena memiliki kualitas *audio* yang jernih, *sharing video* dengan kualitas HD dan merupakan aplikasi yang direkomendasikan untuk melakukan *meeting* dan kelas *online*. Pengguna menyarankan pihak *zoom* untuk menambahkan tidak hanya opsi sembunyikan peserta *non-video* tetapi juga opsi di mana *host* dapat menyembunyikan *video* peserta atau gambar profil agar tidak muncul.

Asosiasi yang terkait dengan kata "*online*" menunjukkan informasi bahwa *zoom cloud meetings* dinilai baik untuk melakukan aktivitas daring selama masa pandemi ini baik itu kegiatan akademis maupun kegiatan kerja. Pengguna menyarankan *zoom* untuk segera memperbaiki masalah yang ada ketika melakukan kegiatan daring.

Asosiasi kata yang berkaitan dengan "*problem*" menunjukkan informasi bahwa pengguna merasa ketika menjalankan aplikasi tidak mengalami permasalahan yang begitu berarti ketika melakukan *meeting.* Terkadang pengguna juga mengalami permasalahan kecil ketika melakukan *meeting* seperti gangguan jaringan yang mengakibatkan menurunnya kualitas *audio* dan *video.* Selain itu, pengguna meminta permasalahan-permasalahan yang ada pada *zoom cloud meetings* untuk segera diperbaiki dan dicari solusinya.

Asosiasi kata yang berkaitan dengan "*time*" menunjukkan bahwa pengguna merasa puas dengan adanya fitur *real time* pada aplikasi *zoom* dan pengguna merasa senang dapat terhubung dengan temannya saat kelas *online* bahkan pengguna merekomendasikan menggunakan aplikasi *zoom cloud meetings* untuk melakukan kelas *online*. Adapun pengguna memberikan saran agar waktu/durasi *meeting* pada *zoom cloud meetings* ini ditambah lebih dari 40 menit karena 40 menit merupakan waktu yang terlalu sebentar menurut pengguna bahkan pengguna berharap bagi pengguna gratisan juga dapat menikmati waktu *meeting* yang *unlimited* seperti aplikasi *video conference* lainnya.

Asosiasi kata yang berkaitan dengan "*work*" menunjukkan informasi bahwa pengguna menilai aplikasi ini berjalan dengan sangat baik ketika dioperasikan dan aplikasi ini dinilai baik dan banyak digunakan untuk melakukan *meeting* kerja.

Asosiasi kata yang berkaitan dengan kata "*connect*" menunjukkan informasi bahwa pengguna menilai untuk konek ke *meeting* di aplikasi *zoom cloud meetings* ini terbilang sangat mudah, simpel dan cepat untuk terhubung. Untuk masalah koneksi pengguna tidak mengalami masalah yang begitu berarti dan sebagian pengguna masih ada yang merasakan masalah *disconnect*. Pengguna menyarankan agar ada fitur notifikasi untuk *connect* ke *meeting* yang telah dijadwalkan dan akan dihadiri dan berharap fitur ini dapat diterapkan bagi pengguna PC ataupun *smartphone*.

Asosiasi kata yang berkaitan dengan kata "*audio*" menunjukkan bahwa pengguna merasa puas dengan kualitas *audio* yang dimiliki oleh *zoom* pada saat melakukan *video meeting*. Beberapa pengguna tidak mengalami masalah pada *audio* dan beberapa pengguna juga memberikan saran agar masalah-masalah pada *audio* untuk segera diperbaiki.

Asosiasi kata yang berkaitan dengan kata "*update*" menunjukkan bahwa pengguna merasa senang dan puas terhadap *update* terbaru dari *zoom*. Pengguna juga mengharapkan untuk *update* berikutnya *zoom* menciptakan fitur baru yang menarik.

b. Ulasan Negatif

Problem		Meeting		Video		Class		Work	
solve	7.211	join	7.004	participant	9.116	online	13.578	doesn't	8.286
facing	6.909	able	3.895	quality	5.809	connect	2.443	meeting	1.986
fix	4.567	connect	3.442	audio	3.926	problem	2.365	problem	1.780
network	4.222	times	3.426	update	2.903	time	2.316		
lot	3.055	error	3.298	poor	2.371	meeting	1.698		
sometime	2.675	network	2.786	problem	2.224				
update	2.644	fix	2.526	meeting	1.639				
hear	2.603	time	2.512						
audio	2.595	problem	2.106						
connect	2.556	work	1.986						
voice	2.500	update	1.809						
one	2.490	class	1.698						
sound	2.445	video	1.639						
class	2.365	audio	1.562						
video	2.224								
meeting	2.106								
online	1.915								
join	1.820								
time	1.820								
work	1.780								

Connect		Audio		Update		Time		Quality	
internet	8.204	hear	5.310	new	13.592	every	9.883	poor	5.870
error	5.823	quality	5.063	fix	4.903	waste	4.332	video	5.809
network	5.254	video	3.926	audio	3.275	connect	2.980	audio	5.063
meeting	3.442	update	3.275	video	2.903	meeting	2.512	sound	4.690
time	2.980	problem	2.595	problem	2.644	class	2.316		
problem	2.556	meeting	1.562	meeting	1.809	problem	1.820		
class	2.443								

Gambar 4.12 Asosiasi Kata Ulasan Negatif

Berdasarkan Gambar 4.12 diperoleh asosiasi kata pada ulasan negatif. Proses ekstraksi informasi dengan asosiasi dilakukan secara berulang kali dengan cara menyaring kata-kata yang memiliki hubungan dengan kata lain dan didasarkan pada relevansi kata dengan topik yang diulas.

Berdasarkan asosiasi yang terkait dengan kata "*meeting*" diperoleh informasi bahwa pengguna beberapa kali gagal untuk *join meeting* padahal koneksi pengguna bagus. Pada saat *meeting* terkadang terjadi *disconnect* secara tiba-tiba dan ini mengganggu kegiatan daring pengguna, pengguna juga mengeluhkan suara yang putus-putus dan *video* nge*blur* saat *meeting* yang kemungkinan disebabkan oleh faktor gangguan jaringan, gangguan jaringan juga dapat disebabkan karena faktor cuaca. Maka dari itu, pengguna mengharapkan akan perbaikan segera untuk permasalahan yang terjadi saat melakukan *video meeting*.

Kata-kata yang berasosiasi dengan kata "*problem*" menunjukkan informasi bahwa pengguna mengalami permasalahan seperti masalah jaringan, *audio*, *video*, dll yang dimana permasalahan ini mengganggu kenyamanan pengguna ketika melakukan aktivitas daring seperti kelas *online*, kerja, dll. Pengguna berharap untuk *update* berikutnya masalah yang dialami pengguna dapat segera diatasi dan dicari solusinya secepat mungkin.

Kata-kata yang berasosiasi dengan kata "*video*" menunjukkan informasi bahwa setelah melakukan *update* aplikasi pengguna mengalami masalah mengenai *video* yang tidak dapat ditampilkan ke partisipan lain padahal *video* sudah dihidupkan. Sejumlah pengguna menilai kualitas *video* yang ditampilkan tidak jelas/*blur* dan tidak bagus.

Kata-kata yang berasosiasi dengan kata "*class*" menunjukkan informasi bahwa pengguna tidak dapat menghadiri kelas *online* dikarenakan permasalahan koneksi

yang kurang mendukung, hal ini dikarenakan pengguna yang bertempat tinggal di desa. Pengguna berharap untuk *update* selanjutnya, internet dengan *low bandwidth* dapat *join meeting* saat kelas online supaya pengguna yang berstatus sebagai pelajar tidak ketinggalan pelajaran.

Kata-kata yang berasosiasi dengan "*work*" menunjukkan informasi bahwa pengguna merasa setelah melakukan *update*, aplikasi justru tidak dapat dijalankan bahkan ada beberapa pengguna mengeluh akan aplikasi yang telah dilakukan *update* justru ketika *meeting* mengalami masalah *audio* yang tidak dapat didengar peserta lain dan *video* yang tidak muncul di layar peserta lain.

Kata-kata yang berasosiasi dengan "*connect*" menunjukkan informasi bahwa pengguna mengeluh karena tidak dapat mengikuti aktivitas secara daring dikarenakan jaringan koneksi internet yang kurang bagus, tetapi ada beberapa pengguna yang mengeluh karena tidak bisa *join meeting* padahal koneksi internetnya dalam kondisi bagus. Pengguna juga mengeluhkan akan koneksi yang kurang bagus akan mempengaruhi kualitas *video* dan *audio.*

Kata-kata yang berasosiasi dengan "*audio*" menunjukkan informasi bahwa kualitas *audio* ketika melakukan *meeting* tidak jernih dan masih terdapat *noise*. Peserta lain terkadang tidak dapat mendengar apa yang dibicarakan oleh *speaker* karena suara yang kurang jelas dan putus-putus, hal ini bisa jadi dikarenakan gangguan jaringan, cuaca,dll.

Kata-kata yang berasosiasi dengan kata "*update*" menunjukkan informasi bahwa pengguna menilai *update* terbaru dari *zoom* justru memakan banyak data *storage* di HP. Pengguna juga merasa dengan melakukan *update* aplikasi justru timbul masalah baru seperti masalah *audio* yang tidak muncul ketika melakukan *screen share*. Pengguna juga menyarankan agar *update* berikutnya merupakan *update* yang efektif seperti tidak memakan banyak *storage* HP dan memperbaiki *bug* dan masalah yang sering terjadi baik pengguna PC ataupun HP.

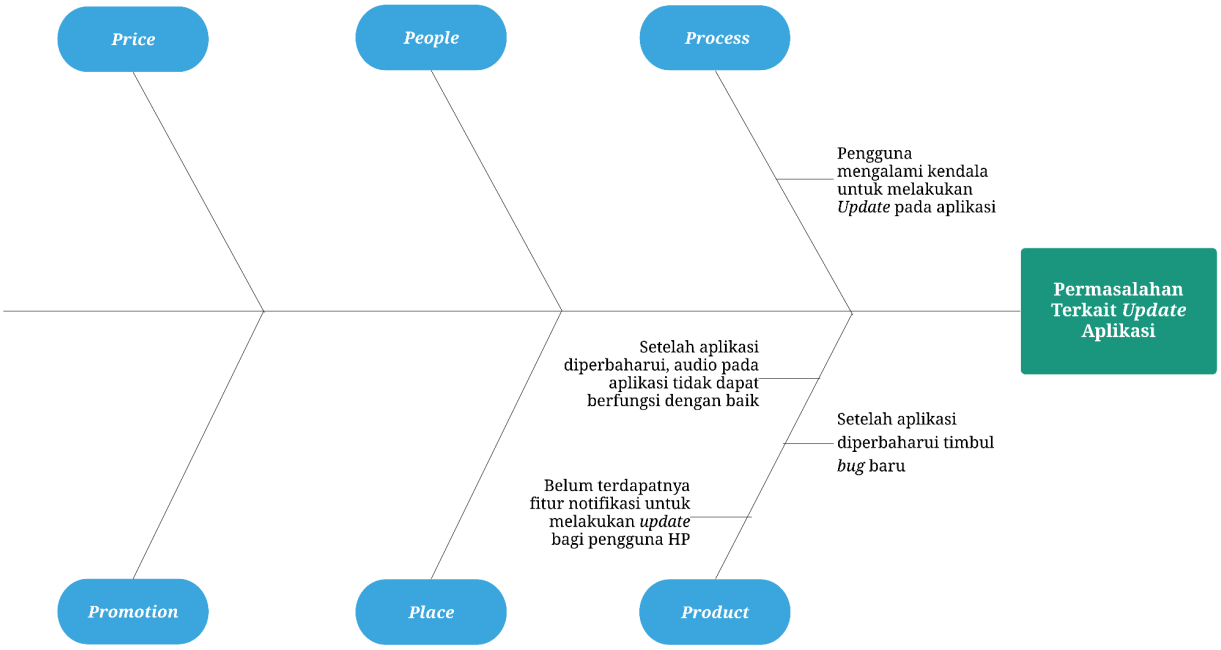
Kata-kata yang berasosiasi dengan kata "*time*" menunjukkan bahwa pengguna mengeluhkan akan keterbatasan waktu yang diterapkan *zoom* untuk melakukan *video meeting* khususnya bagi pengguna gratis, beberapa pengguna juga mengalami masalah *connecting* ke *meeting* dalam durasi yang lama dimana hal ini menurut pengguna merupakan buang-buang waktu saja hanya untuk menunggu *connecting*.

Pengguna menyarankan agar durasi *meeting* bagi pengguna gratis tidak ada batasan/*unlimited* seperti aplikasi *video conference* lainnya.

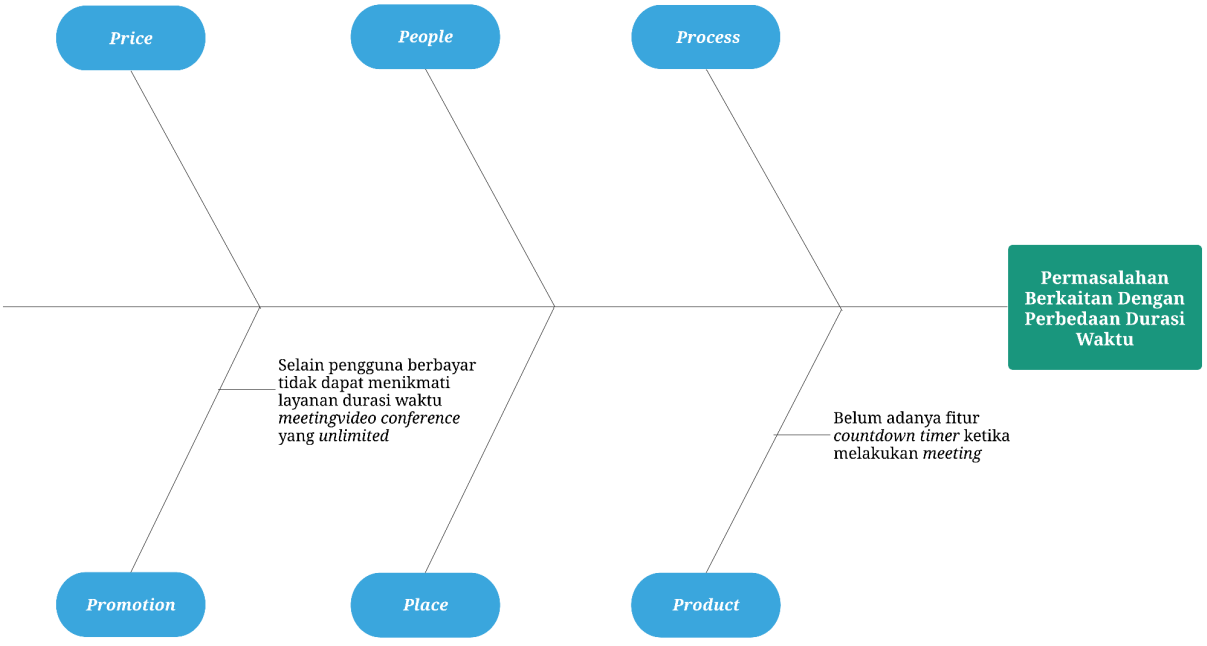
Kata-kata yang berasosiasi dengan kata "*quality*" menunjukkan bahwa pengguna merasa kurang puas akan kualitas yang dimiliki oleh *zoom*, karena kualitas suara dan *video* yang dianggap kurang bagus dan kurang jernih ketika melakukan meeting. Pengguna sangat menyarankan *zoom* untuk segera memperbaiki masalah-masalah seperti ini dan meningkatkan kualitas aplikasi khususnya pada fitur *video* dan *audio*..

4.2.7 *Fishbone Diagram*

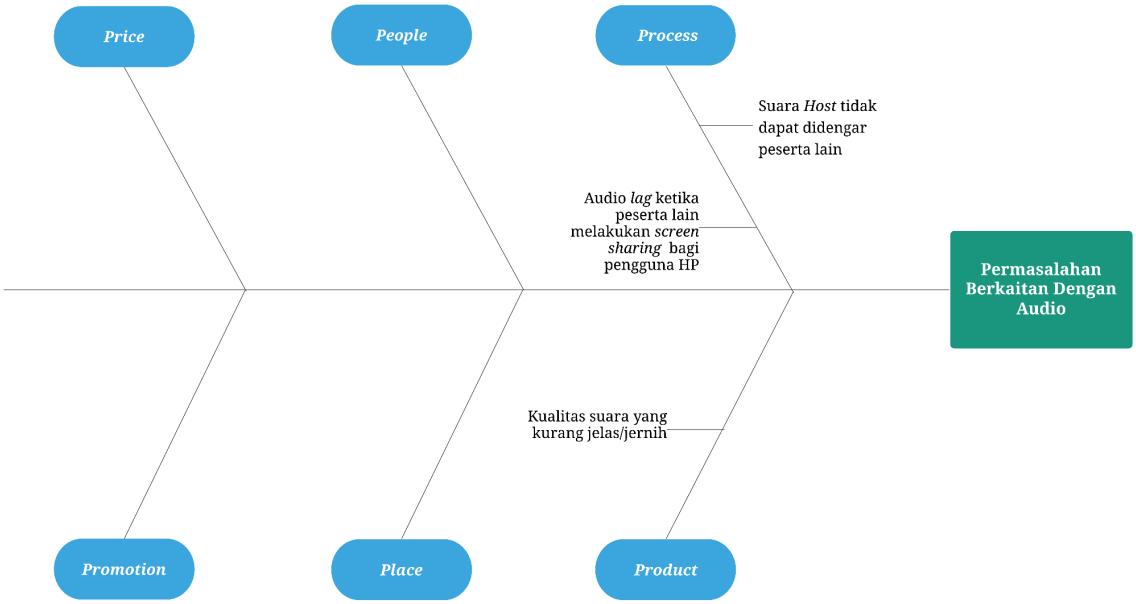
Berdasarkan hasil ulasan negatif yang didapatkan dari data, maka dapat diperoleh informasi mengenai masalah yang terjadi terkait dengan layanan aplikasi *zoom cloud meetings* dan akar penyebab masalahnya yang dipresentasikan pada diagram *fishbone* berikut:



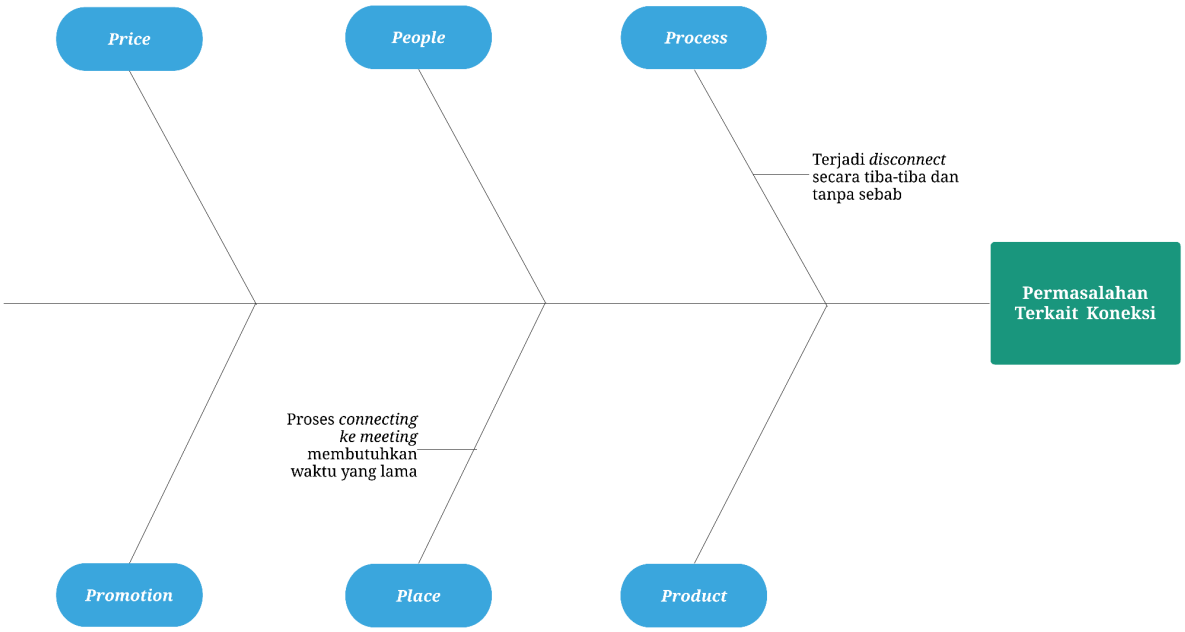
Gambar 4.13 Diagram *Fishbone* Untuk Permasalahan *Update*



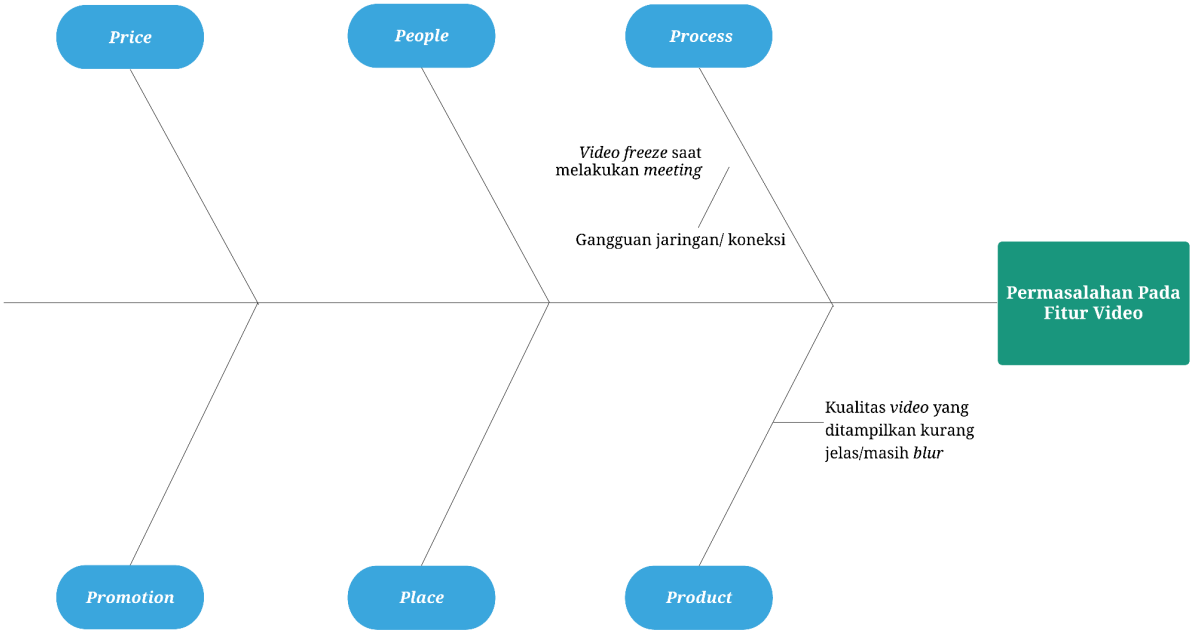
Gambar 4.14 Diagram *Fishbone* Untuk Permasalahan Durasi



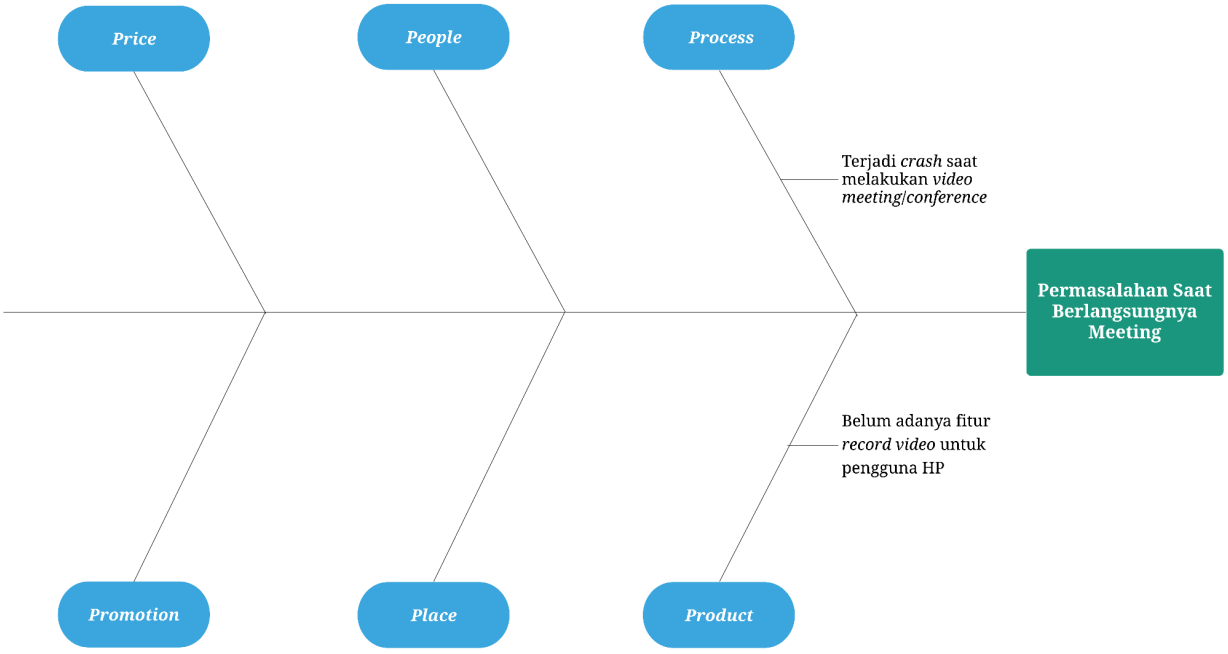
Gambar 4.15 Diagram *Fishbone* Untuk Permasalahan Audio



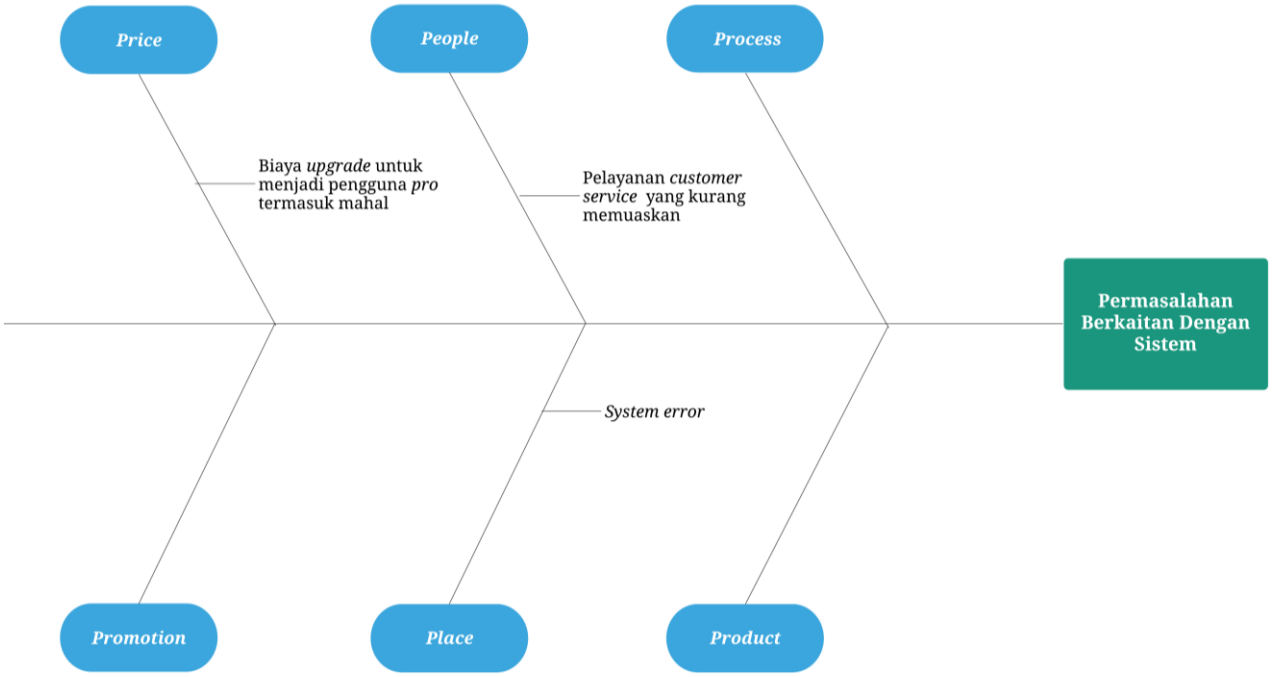
Gambar 4.16 Diagram *Fishbone* Untuk Permasalahan Koneksi



Gambar 4.17 Diagram *Fishbone* Untuk Permasalahan Fitur Video



Gambar 4.18 Diagram *Fishbone* Untuk Permasalahan *Meeting*



Gambar 4.19 Diagram *Fishbone* Untuk Permasalahan Sistem

Berdasarkan analisa dengan diagram *fishbone* dapat diketahui bahwa untuk setiap kata kunci yang terdapat dalam ulasan negatif merupakan penyebab permasalahan selama penggunaan aplikasi. Permasalahan yang terdapat pada analisa diagram *fishbone* didasarkan pada informasi yang didapatkan pada tahap klasifikasi, visualisasi dan asosiasi yang telah dilakukan sebelumnya.

BAB V

ANALISIS DAN PEMBAHASAN

5.1 Gambaran Umum Persepsi Pengguna *Zoom Cloud Meetings*

Berdasarkan data ulasan dari pengguna *zoom* pada 14 Maret 2019 sampai dengan 20 November 2020, didapatkan ulasan sebanyak 96.122. Pada kurun waktu tersebut diketahui bahwa ulasan terbanyak ada pada bulan Juli 2020 seperti yang dapat dilihat pada Gambar 4.8 sebelumnya. Hal ini terjadi dikarenakan pada bulan Maret-Juli terjadi lonjakan kenaikan jumlah unduhan aplikasi *zoom* dikarenakan setiap negara memberlakukan sistem *physical distancing* dan *Work From Home* (WFH) untuk mencegah penularan virus corona selama masa pandemi ini sehingga penggunaan aplikasi layanan yang mendukung fitur *video conference* termasuk *zoom cloud meetings* juga turut meningkat. Kemudian dari aspek *rating*, sebanyak 29.95% pengguna atau 28.791 pengguna memberikan *rating* 1 atau tidak bagus. Setelah dilakukan analisis deskriptif untuk melihat gambaran umum dari pengguna *zoom* selanjutnya dilakukan *preprocessing* data. Pada tahap *preprocessing* data ulasan yang awalnya berjumlah 96.122 kemudian berubah jumlahnya menjadi 95.177. ulasan. Banyaknya jumlah ulasan yang berkurang ini disebabkan karena adanya ulasan yang hanya berupa *emoticon* atau simbol, angka dan spasi sehingga data ini dianggap tidak penting dan dihilangkan ketika dilakukan pembersihan data.

Pada ulasan yang mengandung pernyataan positif seperti ungkapan memuji, ungkapan rasa terimakasih dan lain sebagainya dimasukkan ke dalam kategori sentimen positif. Adapun untuk ulasan klasifikasi yang mengandung pernyataan negatif seperti penghinaan, ketidakpuasan, dan sebagainya dimasukkan ke dalam kategori sentimen negatif. Oleh karena itu hanya digunakan dua pelabelan yaitu sentimen positif dan negatif. Berdasarkan hasil akhir pelabelan diperoleh informasi bahwa jumlah ulasan sentimen positif memiliki frekuensi lebih besar dibanding sentimen negatif. Hal ini ditunjukkan bahwa, dimana ulasan pada kelas sentimen positif berjumlah 73,44% atau 55.413 ulasan. Sedangkan 26,56% atau 20.041 ulasan termasuk sentimen negatif.

5.2 Hasil Klasifikasi Metode *Naïve Bayes Classifier*

Tahap awal klasifikasi yaitu dengan membagi jumlah data uji dan data latih. Pada penelitian ini dilakukan perbandingan dengan persentase data latih dan data uji yang berbeda sebanyak lima kali. Pembagian data pertama diawali dengan data latih 60% (45.273 ulasan) dan data uji 40% (30.181 ulasan). Pada pembagian data latih dan data uji yang pertama diketahui *Class Precision* negatif 92,35% sedangkan untuk kelas positif sebesar 83,26%. Akurasi yang didapatkan adalah sebesar 86,31%.

Pembagian data kedua dilakukan dengan data latih 70% (52.818 ulasan) dan data uji 30% (22.636 ulasan). Pada pembagian data latih dan data uji yang kedua dihasilkan *Class Precision* kelas negatif sebesar 91,83%sedangkan untuk kelas positif sebesar 83,23% dengan akurasi yang didapatkan lebih tinggi dibandingkan pembagian data pertama, yaitu 86,15%.

Pembagian data ketiga dilakukan dengan data latih 80% (60.363 ulasan) dan data uji 20% (15.091 ulasan). Pada pembagian data latih dan data uji yang ketiga dihasilkan *Class Precision* kelas negatif sebesar 91,45% sedangkan untuk kelas positif meningkat menjadi 83,57%, dengan akurasi yang dihasilkan lebih tinggi dibandingkan pembagian data pertama dan kedua, yaitu 86,30%.

Pembagian data keempat dilakukan dengan data latih 85% (64.136 ulasan) dan data uji 15% (11.318 ulasan). Pada pembagian data latih dan data uji yang keempat dihasilkan *Class Precision* kelas negatif sebesar 91,54% sedangkan untuk kelas positif meningkat menjadi 84,08%, dengan akurasi yang dihasilkan lebih tinggi dibandingkan tiga perbandingan data sebelumnya, yaitu 86,67%.

5.3 Hasil *Association Rules*

Berdasarkan visualisasi yang telah dilakukan pada ulasan positif di bab sebelumnya dengan jumlah ulasan sebanyak 55.413 ulasan, didapatkan beberapa kata yang paling sering muncul, yang kemudian dikerucutkan menjadi 10 kata dengan frekuensi terbanyak antara lain : kata "*meeting*" sebanyak 5441 kali, kata "*class*" sebanyak 2878 kali, kata "*video*" sebanyak 2667 kali, kata "*online*" sebanyak 2494 kali, kata "*time*" sebanyak 1483 kali kata "*problem*" sebanyak 1442 kali, kata "*work*" sebanyak 1386 kali, kata "*connect*" sebanyak 1140 kali, kata "*audio*" sebanyak 985 kali, dan kata "*update*" sebanyak 945 kali.

Asosiasi kata pada ulasan positif dilakukan dengan melakukan ekstraksi informasi secara berulang-ulang dan didasarkan pada relevansi kata dengan topik yang terkait. Pada Gambar 4.13 menunjukkan bahwa asosiasi kata dilakukan pada kata "*meeting*", kata "*class*", kata "*video*", kata "*online*", kata "*problem*", kata "*time*", kata "*work*", kata "*connect*", kata "*audio*", dan kata "*update*".

Berdasarkan tahap *association rule* yang telah dilakukan diketahui bahwa dari segi kegiatan *meeting* aplikasi *zoom cloud meetings* telah memiliki performa yang baik, kualitas yang bagus baik itu dari *audio, video, screen,*dll dan juga *zoom* dinilai baik ketika kita menjadi *host* ataupun peserta *meeting* sehingga pengguna merekomendasi aplikasi ini untuk melakukan *meeting*. Aplikasi *zoom cloud meetings* ini dinilai baik oleh pelajar karena dapat menghadiri kelas secara daring khususnya di masa pandemi ini. Pengguna memberikan saran agar durasi waktu untuk melakukan *video* tidak dibatasi bagi pengguna gratis. Menurut pengguna aplikasi *zoom cloud meetings* ini merupakan aplikasi *video conference* terbaik karena memiliki kualitas *audio* yang jernih, kualitas *sharing video* dengan kualitas HD dan merupakan aplikasi yang direkomendasikan untuk melakukan *meeting* dan kelas *online*. Pengguna menyarankan pihak *zoom* untuk menambahkan tidak hanya opsi sembunyikan peserta *non-video* tetapi juga opsi di mana *host* dapat menyembunyikan *video* peserta atau gambar profil agar tidak muncul. Pengguna merasa ketika menjalankan aplikasi tidak mengalami permasalahan yang begitu berarti ketika melakukan *meeting.* Terkadang pengguna juga mengalami permasalahan kecil ketika melakukan *meeting* seperti gangguan jaringan yang mengakibatkan menurunnya kualitas *audio* dan *video*. Selain itu, pengguna meminta permasalahan-permasalahan yang ada pada *zoom cloud meetings* untuk segera diperbaiki dan dicari solusinya. Pengguna merasa puas dengan adanya fitur *realtime* pada aplikasi *zoom* dan pengguna merasa senang dapat terhubung dengan temannya saat kelas *online* bahkan pengguna merekomendasikan menggunakan aplikasi *zoom cloud meetings* untuk melakukan kelas *online.* Adapun pengguna memberikan saran agar waktu/durasi *meeting* pada *zoom cloud meetings* ini ditambah lebih dari 40 menit karena 40 menit merupakan waktu yang terlalu sebentar menurut pengguna bahkan pengguna berharap bagi pengguna gratisan juga dapat menikmati waktu *meeting* yang *unlimited*

seperti aplikasi *video conference* lainnya. Pengguna menilai aplikasi ini berjalan dengan sangat baik ketika dioperasikan dan aplikasi ini dinilai baik dan banyak digunakan untuk melakukan *meeting* kerja. Pengguna menilai untuk terhubung ke *meeting* di aplikasi *zoom cloud meetings* ini terbilang sangat mudah, simpel dan cepat untuk terhubung. Untuk masalah koneksi pengguna tidak mengalami masalah yang begitu berarti dan sebagian pengguna masih ada yang merasakan masalah *disconnect*. Pengguna menyarankan agar ada fitur notifikasi untuk *connect* ke *meeting* yang telah dijadwalkan dan akan dihadiri dan berharap fitur ini dapat diterapkan bagi pengguna PC ataupun *smartphone*. Pengguna merasa puas dengan kualitas *audio* yang dimiliki oleh *zoom* pada saat melakukan *video meeting*. Beberapa pengguna tidak mengalami masalah pada *audio* dan beberapa pengguna juga memberikan saran agar masalah-masalah pada *audio* untuk segera diperbaiki.

Berdasarkan 20.041 ulasan negatif diketahui bahwa terdapat 10 kata yang paling banyak muncul seperti yang dapat dilihat pada Gambar 4.10. Kata yang paling banyak muncul diantaranya: kata "*problem*" sebanyak 3030 kali kata, "*meeting*" sebanyak 2369 kali, kata "*video*" sebanyak 1663 kali, kata "*audio*" sebanyak 1328 kali, kata "*connect*" sebanyak 1272 kali, kata "*class*" sebanyak 1250 kali, kata "*time*" sebanyak 1066 kali, kata "*update*" sebanyak 1000 kali kata "*quality*" sebanyak 991 kali dan kata "*work*" sebanyak 981 kali.

Asosiasi kata pada ulasan negatif dilakukan dengan melakukan ekstraksi informasi secara berulang-ulang dan didasarkan pada relevansi kata dengan topik yang terkait. Pada Gambar 4.14 didapatkan informasi bahwa asosiasi kata dilakukan pada kata "*meeting*", kata "*problem*", kata "*video*", kata "*class*", kata "*work*", kata "*connect*", kata "*audio*", kata "*quality*", kata "*time*", dan kata "*online*".

Pengguna mengalami beberapa kali gagal untuk *join meeting* padahal koneksi pengguna bagus. Pada saat *meeting* terkadang terjadi *disconnect* secara tiba-tiba dan ini mengganggu kegiatan daring pengguna, Maka dari itu, pengguna mengharapkan akan perbaikan segera untuk permasalahan yang terjadi saat melakukan *video meeting*. Pengguna mengalami permasalahan seperti masalah jaringan, *audio, video,* dll yang dimana permasalahan ini mengganggu kenyamanan pengguna ketika melakukan aktivitas daring seperti kelas *online,* kerja, dll. Pengguna berharap untuk *update* berikutnya masalah yang dialami pengguna dapat segera diatasi dan dicari solusinya secepat mungkin. Setelah aplikasi dilakukan *update* pengguna mengalami masalah

mengenai *video* yang tidak dapat ditampilkan ke partisipan lain padahal *video* sudah dihidupkan. Sejumlah pengguna menilai kualitas *video* yang ditampilkan tidak jelas/*blur* dan tidak bagus. Pengguna tidak dapat menghadiri kelas *online* dikarenakan permasalahan koneksi yang kurang mendukung, hal ini dikarenakan pengguna yang bertempat tinggal di desa. Pengguna berharap untuk *update* selanjutnya, internet dengan *low bandwidth* dapat *join meeting* saat kelas online supaya pengguna yang berstatus sebagai pelajar tidak ketinggalan pelajaran. Pengguna merasa aplikasi setelah dilakukan *update* justru tidak dapat dijalankan, bahkan ada beberapa pengguna mengeluh akan aplikasi yang sudah dilakukan *update* justru ketika *meeting* mengalami masalah *audio* yang tidak dapat didengar peserta lain dan *video* yang tidak muncul di layar peserta lain. Pengguna mengeluh karena tidak dapat mengikuti aktivitas secara daring dikarenakan jaringan koneksi internet yang kurang bagus, tetapi ada beberapa pengguna yang mengeluh karena tidak bisa *join meeting* padahal koneksi internetnya dalam kondisi bagus. Pengguna juga mengeluhkan akan koneksi yang kurang bagus akan mempengaruhi kualitas *video* dan *audio.*

Pengguna menilai *update* terbaru dari *zoom* justru memakan banyak data *storage* di HP. Pengguna juga merasa dengan melakukan *update* aplikasi justru timbul masalah baru seperti masalah *audio* yang tidak muncul ketika melakukan *screen share*. Pengguna juga menyarankan agar *update* berikutnya merupakan *update* yang efektif seperti tidak memakan banyak *storage* HP dan memperbaiki *bug* dan masalah yang sering terjadi baik pengguna PC ataupun HP. Kualitas *audio* ketika melakukan *meeting* tidak jernih dan masih terdapat *noise.* Peserta lain terkadang tidak dapat mendengar apa yang dibicarakan oleh *speaker* karena suara yang kurang jelas dan putus-putus, hal ini bisa jadi dikarenakan gangguan jaringan, cuaca,dll. Pengguna mengeluhkan akan keterbatasan waktu yang diterapkan *zoom* untuk melakukan *video meeting* khususnya bagi pengguna gratis, beberapa pengguna juga mengalami masalah *connecting* ke *meeting* dalam durasi yang lama dimana hal ini menurut pengguna merupakan buang-buang waktu saja hanya untuk menunggu *connecting*. Pengguna menyarankan agar durasi *meeting* bagi pengguna gratis tidak ada batasan/*unlimited* seperti aplikasi *video conference* lainnya. Pada saat aktivitas *meeting* dan kelas *online*, pengguna mengalami permasalahan seperti *crash,* tidak bisa *join meeting*, dll sehingga mengakibatkan pengguna tidak dapat berpartisipasi dalam aktivitas kelas *online.* Pengguna sangat menyarankan *zoom* untuk segera memperbaiki masalah-masalah seperti ini.

5.4 Rencana Perbaikan Berdasarkan Diagram *Fishbone*

Berdasarkan diagram *fishbone* pada bab sebelumnya diketahui informasi yang menjadi faktor permasalahan pada penggunaan aplikasi *zoom cloud meetings* dan penyebab *zoom* mendapatkan ulasan negatif dan *rating* buruk dari penggunanya. Terdapat 6 aspek yang digunakan yaitu *product, process, people, place, price* dan *promotion*. Langkah selanjutnya yaitu menentukan rencana perbaikan yang didasarkan pada permasalahan yang ada dan berdasarkan saran dari beberapa pengguna. Adapun rencana perbaikannya juga bersifat umum dan tidak menjurus pada hal yang bersifat teknis dan dapat dilihat rencana perbaikan masalah pada tabel berikut :

Tabel 5.1 Rencana Perbaikan Masalah

Aspek	Permasalahan yang Terjadi	Saran Perbaikan
Product	• Kualitas *video* yang ditampilkan kurang jelas/masih *blur* • Kualitas suara yang kurang jelas/jernih • Setelah aplikasi diperbaharui, audio pada aplikasi tidak dapat berfungsi dengan baik • Setelah aplikasi diperbaharui timbul *bug* baru • Belum adanya fitur *record video* untuk pengguna HP • Belum terdapatnya fitur notifikasi untuk melakukan *update* bagi pengguna HP • Belum terdapatnya fitur *countdown timer* ketika melakukan *meeting*	1. *Maintenance* aplikasi untuk memperbaiki *bug* 2. Melakukan penambahan fitur notifikasi *update* bagi pengguna HP 3. Memperbaiki kualitas suara yang masuk agar meminimalisir *noise* 4. Menambah fitur *record video* untuk pengguna HP 5. Memberikan fitur *option* untuk menampilkan kualitas *video* seperti 360p atau kualitas HD 720p
Process	• Pengguna mengalami kendala untuk melakukan *update* pada aplikasi • Terjadi *crash* saat melakukan *video meeting*/*conference* • *Video freeze* saat melakukan *meeting* • Terjadi *disconnect* secara tiba-tiba dan tanpa sebab • Suara *host* tidak dapat didengar oleh peserta lain • Audio *lag* ketika peserta lain melakukan *screen sharing* bagi pengguna HP	1. Memperbaiki *bug* terkait koneksi 2. Melakukan *update* untuk memperbaharui kualitas *video* 3. Melakukan *update* untuk memperbaharui kualitas *audio* dan memperbaiki *bug* pada sistem *audio* 4. Meluncurkan versi terbaru guna menghilangkan *bug* yang terdapat pada versi sebelumnya dan memperbaiki masalah *error*

Aspek	Permasalahan yang Terjadi	Saran Perbaikan
People	• Pelayanan *customer service* yang kurang memuaskan	1. Melakukan evaluasi kinerja pada bagian *customer service* 2. Bagian pelayanan disarankan agar lebih responsif terhadap tanggapan penggunanya terutama tanggapan yang kritis dan pihak perusahaan dapat cepat memperbaiki kekurangan yang ada pada aplikasinya
Place	• *System error* • Proses *connecting* ke *meeting* membutuhkan waktu yang lama	1. Sistem *error* umumnya terjadi ketika ada *update* pada aplikasi. Maka dari itu, sebelum meluncurkan versi aplikasi terbaru pihak *developer* lebih memperhatikan dalam tahap pengujian versi hingga aplikasi tidak terdeteksi *error* 2. *Developer* segera memperbaiki server pada aplikasi sehingga tidak terjadi masalah terkait *connecting error* atau *connecting buffering*
Price	• Biaya *upgrade* untuk menjadi pengguna *pro* tergolong mahal	1. Mempertimbangkan kembali terhadap biaya *upgrade* dan memberikan potongan harga *upgrade* bagi penggunanya
Promotion	• Selain pengguna berbayar tidak dapat menikmati durasi *meeting* yang *unlimited*	1. Mempertimbangkan kembali untuk menerapkan kebijakan bagi pengguna gratis untuk dapat menikmati durasi *meeting* yang *unlimited* atau durasi lebih dari 60 menit

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan pada bab sebelumnya, maka dapat ditarik beberapa kesimpulan seperti berikut:

1. Berdasarkan ulasan yang telah dikumpulkan dapat diketahui bahwa mayoritas pengguna *Zoom Cloud Meetings* mempunyai penilaian ataupun persepsi yang baik terhadap aplikasi tersebut. Pada kurun waktu 14 Maret 2019 – 20 November 2020 didapatkan ulasan pengguna *zoom* pada laman web *Google Play* yaitu sebanyak 96.122 ulasan. Adapun pengguna paling banyak memberikan rating 5 dengan persentase 48.99% atau sebanyak 47.086 pengguna. Setelah dilakukan *preprocessing* diketahui ulasan yang dapat diolah sebanyak 1.843 yang kemudian diperoleh jumlah ulasan positif sebanyak 62.896 ulasan (65,71%) dan ulasan negatif sebanyak 32.821 ulasan (34,29%).
2. Berdasarkan klasifikasi yang telah dilakukan dengan metode *Naïve Bayes Classifier*, yang dimana dengan menggunakan empat jenis perbandingan data latih dan data uji. Diketahui bahwa dengan data latih dan data uji sebesar 85% : 15% menghasilkan tingkat akurasi sebesar 86,67%.
3. Berdasarkan hasil asosiasi yang dilakukan, didapatkan informasi bahwa secara umum penelitian ini membahas delapan hal yakni *meeting, video, audio, class, work, update, time* dan *connect*. Pada asosiasi positif pengguna banyak membahas mengenai kata“*meeting*”, “*class*”, “*video*”, “*online*”, “*problem*”, “*time*”, “*work*”, “*connect*”, “*audio*”, dan “*quality*”. Sedangkan pada asosiasi negatif pengguna sering membicarakan mengenai kata “*meeting*”, “*problem*”, “*video*”, “*class*”, “*work*”, “*connect*”, “*audio*”, “*update*”, “*time*”, “*online*”.
4. Berdasarkan diagram *fishbone* didapatkan permasalahan yang didasarkan pada 6 faktor yaitu *product, process, people, place, price* dan *promotion*. Dari keenam faktor tersebut diperoleh permasalahan sebanyak 18, antara lain : terjadi *crash* saat melakukan *video meeting/conference,* belum adanya fitur *record video* untuk

pengguna HP, *video freeze* saat melakukan *meeting***,** *video* yang ditampilkan kurang jelas/*blur*, suara *host* tidak dapat didengar oleh peserta lain, kualitas suara yang kurang jernih, terjadi *disconnect* secara tiba-tiba dan tanpa sebab**,** *system error*, munculnya *bug* setelah *update* aplikasi, aplikasi tidak dapat dilakukan *update*, belum ada notifikasi *update* bagi pengguna HP, pelayanan yang kurang memuaskan, biaya *upgrade* untuk menjadi pengguna *pro* termasuk mahal, selain pengguna berbayar tidak dapat menikmati durasi *meeting* yang *unlimited*, dan proses *connecting* ke *meeting* dalam waktu yang lama.

6.2 Saran

Berdasarkan beberapa kesimpulan diatas, adapun saran yang dapat diberikan adalah sebagai berikut :

1. Saran bagi pihak *Zoom*

   Bagi pihak *Zoom*, hasil ekstraksi informasi dari ulasan-ulasan yang telah diberikan oleh pengguna khususnya ulasan yang berbentuk negatif dapat dijadikan bahan evaluasi kedepannya dalam hal peningkatan kepuasan pengguna dan memberikan pelayanan semaksimal mungkin, serta untuk pengembangan pembaharuan aplikasi dan fitur selanjutnya.

2. Saran untuk penelitian selanjutnya

   Peneliti selanjutnya dapat mengembangkan penelitian dengan menggunakan lebih dari satu bahasa dikarenakan penelitian ini masih dibatasi dengan hanya mengolah ulasan berbahasa Inggris. Kemudian penelitian ini baru menggunakan satu algoritma yaitu apriori, sehingga penelitian selanjutnya dapat dilakukan menggunakan pendekatan *association rule* lainnya sebagai pembanding performa algoritma apriori untuk menemukan asosiasi antar kata yang lebih akurat. Selain itu pada penelitian ini hanya menganalisis satu aplikasi *video conference*, untuk penelitian selanjutnya dapat membandingkan lebih dari satu aplikasi. Untuk penelitian selanjutnya diharapkan dapat menerapkan metode lain atau teknik lain terhadap kata negasi agar hasil yang didapatkan lebih akurat.

DAFTAR PUSTAKA

Adiyana, I., & Hakim, F. (2015). Implementasi Text mining pada Mesin Pencarian Twitter untuk Menganalisis Topik-Topik terkait KPK dan Jokowi. *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika UMS*, (pp. 570-581).

APJII. (2020). Laporan Survei Internet APJII 2019 - 2020 [Q2]. (*online*): https://apjii.or.id/survei2019x/download/yIzbeAuEsRU28Wm34BnTwtxGoqJ51K (16 Januari 2021)

Appfollow. (2020). App Review Management & ASO Platform. (*online*): https://appfollow.io/ (05 Januari 2021)

Aprianti, W., Hafizd, K., & Rizani, M. (2017). Implementasi Association Rules dengan Algoritma Apriori pada Dataset Kemiskinan. *J. Math. and Its Appl., XIV*(2), 145-155.

Arifwidodo, B., & Ikhwan, S. (2019). Analisa Quality Of Service Layanan Video Call Berbasis Internet Protocol Multi Media Subsystem Pada Jaringan Ip Versi 6. *Jurnal Nasional, 20*(2), 95-104.

Arkhamsiagustinah. (2015). Perbandingan Metode Second-Order Fuzzy Time Series Dari HSU dan CHEN Dalam Peramalan Produk Domestik Regional Bruto (PDRB). *Skripsi*.

Azizah, N. (2019). Pengambilan Keputusan Seleksi Penerima Bantuan Rumah Tidak Layak Huni (Rtlh) Penduduk Desa Kalianyar Dengan Menggunakan Algoritma Naive Bayes. *Skripsi*, Universitas Pelita Bangsa.

Carracedo, P., Puertas, R., & Marti, L. (2020). Research Lines On The Impact Of The COVID-19 Pandemic On Business. A Text Mining Analysis. *Journal of Business Research*.

Chintalapudi, N., Battineni, G., Canio, M., Sagaro, G., & Amenta, F. (2020). Text Mining with Sentiment Analysis on Seafarers' Medical Documents. *International Journal of Information Management Data*.

Dunham, M. (2003). *Data Mining: Introductory And Advanced Topics.* New Jersey: Pearson Education.

Fanani. (2017). *Klasifikasi Review Software Pada Google Play Menggunakan Pendekatan*. Skripsi. Universitas Gadjah Mada.

Fauziah, N. (2009). *Aplikasi Fishbone Analysis dalam Meningkatkan Kualitas Produksi Teh Pada PT Rumpun Sari Kemuning, Kabupaten Karanganyar*. Skripsi. Universitas Sebelas Maret.

Fauzy, M., Asror, I., & Wiharja K, R. (2016). Penerapan Metode Association Rule Menggunakan Algoritma Apriori Pada Simulasi Prediksi Hujan Wilayah Kota Bandung. *Jurnal Ilmiah Teknologi Informasi Terapan, II*(2).

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters 27.8*, 861–874.

Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook:Advanced Approaches to Analyzing Unstructured Data.* New York: Cambridge University Press.

Fikria, N. (2018). Analisis Klasifikasi Sentimen Review Aplikasi E-Ticketing Menggunakan Metode Support Vector Machine Dan Asosiasi. *Skripsi*.

GooglePlay. (2020). *details: Zoom Cloud Meetings*. (*online*): https://play.google.com/store/apps/details?id=us.zoom.videomeetings&hl=en&gl=US (31 Desember 2020)

Gough, M. (2006). *Video Conferencing Over IP.* Canada: Syngress Publishing.

Han, J., & Kamber, M. (2006). *Data Mining : Concepts and Techniques Second Edition.* San Fransisco: Morgan Kauffman.

Hansen, & Mowen. (2001). Akuntansi Manajemen Biaya Jilid 2. Jakarta: Erlangga.

Hansen, & Mowen. (2006). Akuntansi Manajemen 1. Jakarta: Salemba Empat.

Herjanto, E. (2007). Manajemen Operasi. Jakarta: PT.Gramedia Widia Sarana.

Herlinawati, N., Yuliani, Y., Faizah, S., Gata, W., & Samudi. (2020). Analisis Sentimen Zoom Cloud Meetings Di Play Store Menggunakan Naïve Bayes dan Support Vector Machine. *Journal of Computer Engineering System and Science, V*(02).

Herryawan, K. (2009). Modul Pelatihan Teknisi JARDIKNAS dan SIM Keuangan. *Video Conference*.

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media.*

Institute, S. (2010). *Getting Started with SAS® Text Miner 4.2.* North Carolina: SAS Publishing.

Interbrand. (2020). *Best Global Brand 2020*. (*online*): https://interbrand.com/newsroom/interbrand-reveals-2020-best-global-brands-report/ (20 Desember 2020)

Josi, A., Abdillah, L., & Suryayusra. (2014). Penerapan Teknik Web Scraping Pada Mesin Pencari Artikel Ilmiah. *Jurnal Sistem Informasi, V*(2), 159-164.

Juliasari, N., & Sitompul, J. C. (2012). Aplikasi Search Engine dengan Metode Depth First Search (DFS). *BIT Numerical Mathematics*, 9.

Karch. (2019). *What is Google Play?* Retrieved from http://lifewire.com/what-isgoogle-play-891616720.html

Karyadi, S., Yasin, H., & Mukid, M. (2016). Analisis Kecenderungan Informasi Dengan Menggunakan Metode Text Mining. *Jurnal Gaussian, V*(4), 763-770.

Kiran, D. (2016). *Total Quality Management 1st Edition: Key Concepts and Case Studies*. Oxford: Butterworth-Heinemann.

Kotler, P., & Amstrong, G. (1996). *Dasar-Dasar Pemasaran* (5th ed.). Jakarta: Intermedia.

Kurniasari, S. (2018). Implementasi SVM dan Asosiasi Untuk Sentiment Analysis Data Ulasan The Phoenix Hotel Yogyakarta Pada Situs Tripadvisor. *Skripsi*.

Liker, J. K. (2014). *Der Toyota Weg* (9.Edition ed.). Munich, Germany: FinanzBuch Verlag.

Liu, L., Chunfang, & Zhouyang. (2013). Analysis of Customer Satisfaction from Chinese Reviews using Opinion Mining. *6th IEEE International Conference on Software Engineering and Service Science (ICSESS)*.

Ma'arif, M. (2016). Integrasi Laman Web Tentang Pariwisata DIY. *Jurnal Informatikan dan komputer, IX*(1).

Manning, C. D., Raghavan, P., & Schütze, H. (2009). *An Introduction to Information Retrieval – Online Edition.* Cambridge: Cambridge University Press.

Mercubuana Jakarta, M. T. (2011). *Metode Quality Function Deployment?? Apa Sich?* (*online*):https://www.kompasiana.com/muslimah/5500daaf813311f51bfa7bb7/metode-quality-function-deployment-apa-sich (28 Desember 2020)

Mu-Hee, S., Soo-Yeon, L., Seong-Bae, P., Dong-Jin, K., & Sang-Jo, L. (2006). Ontology-based automatic classification of web documents. *Applied Soft Computing Technologies: The Challenge of Complexity*, 483-493.

Novienstinawati. (2008). Analisa Implementasi Layanan Video Conference Berbasis Web. *Skripsi*.

Nurdiansyah, D. (2013). *Implementasi Video Conference Pada Jaringan HSUPA (High Speed Uplink Packet Access) Dengan Media IPv6 Menggunakan Simulator OPNET Modeler v.14.5.* Universitas Brawijaya.

Pradana, M. (2020). Penggunaan Fitur Wordcloud dan Document Term Matrix Dalam Text Mining. *Jurnal Ilmiah Informatika, VIII*(1).

Priyanto, A., & Ma'arif, M. (2018). Implementasi Web Scraping dan Text Mining untuk Akuisisi dan Kategorisasi Informasi Laman Web Tentang Hidroponik. *Indonesian Journal of Information Systems (IJIS), I*(1).

Putri, D. (2016). *Implementasi Inferensi Fuzzy Mamdani Untuk Keperluan Sistem Rekomendasi Berita Berbasis Konten.* Skripsi. UGM Yogyakarta: Program Studi Ilmu Komputer FMIPA.

Rozi, I. F., Pramono, S. H., & Dahlan, E. A. (2012). Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi. *Jurnal EECCIS, VI*(1).

SAS Institute. (2010). *Getting Started With SAS® Text Miner 4.2.* North Carolina: SAS Publishing.

Statqo Analytics. (2020). *Press Release (Per 30 Maret 2020 – Edisi ke 2) Analisis Digital Terkait Dampak Pandemi COVID-19.* (*online*): https://www.statqoanalytics.com/post/press-release-per-30-maret-2020-edisi-ke-2-analisis-digital-terkait-dampak-pandemi-covid-19 (15 Desember 2020)

Tan, P., Steinbach, M., & Karpatne, A. (2006). *Introduction To Data Mining.* USA: Addison-Wesley.

Tanaka, M., & Okutomi, M. (2014). A Novel Inference of a Restricted Boltzmann Machine. *22nd International Conference on Pattern Recognition*, (pp. 1526-1531).

Titania, B. (2020). *Penerapan Metode Text Mining Dan Social Network Analysis Pada Jejaring Sosial Twitter*. Skripsi. Universitas Islam Indonesia

Turban, E., & King, D. (2002). *Electronic Commerce 2002 – A Managerial Perspective (Second edition).* New York: Prentice Hall.

Turland, M. (2010). *Php | Architect's Guide to Web Scraping* . Los Angeles: Marco Tabini & Associates, Inc.

Ulwan, M. N. (2016). *Pattern Recognition pada Unstructured Data Teks Menggunakan Support Vector Machine dan Association*. Skripsi. Universitas Gajah Mada. Yogyakarta: Program Studi Statistik FMIPA.

Vaish, R. (2018). *VADER and Sentiment Analysis — Python*. (*online*): https://medium.com/@rahulvaish/vader-and-sentiment-analysis-python-eae70ecef454 (09 Februari 2021)

Valibia, K. G. (2018). *Zoom*. (*online*): UC,Riverside: https://sitelicense.ucr.edu/files/zoom_for_instructors.pdf (10 September 2020)

Wati, R. (2016). Penerapan Algoritma Genetika Untuk Seleksi Fitur Pada Analisis Sentimen Review Jasa Maskapai Penerbangan Menggunakan Naive Bayes. *Evolusi: Jurnal Sains dan Manajemen, IV*(1).

Wijayatun, P. R., & Nugroho. (2016). Prediksi Rating Film Menggunakan Metode Naive Bayes. *Jurnal Teknik Elektro (ISSN 1411-0059)*, 60-63.

Zhang, Y., Chen, M., & Liu, L. (2015). A Review on Text Mining. *International Journal on Software Engineering and Service Science* .

Zoom. (2020). *Zoom Brand Guide*. (*online*): https://zoom.us/docs/doc/Zoom%20Brand%20Guide.pdf (21 Januari 2021)

LAMPIRAN

Lampiran 1. *Script preprocessing*

```
#import library yang dibutuhkan

import numpy as np
import pandas as pd
import  string
import re

df=pd.read_excel('Data Review.xlsx', usecols=['Rating', 'Review'])

df['Rating']=df['Rating'].astype('str')

df1=df.apply(lambda x: x.str.replace( r'[^\w\s]', '').str.lower()).replace(['rating', ''],
np.nan, regex=True)

df1= df1.dropna().reset_index(drop=True)
```

Lampiran 2. *Script* Pelabelan dan Pembobotan

```
#import library yang dibutuhkan

import numpy as np
import pandas as pd
import nltk
import string
import re

from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
analyser =  SentimentIntensityAnalyzer()

df=pd.read_excel('reviews zoom fix.xlsx', usecols=['Rating', 'Review'])
df['Rating']=df['Rating'].astype('str')
df1=df.apply(lambda x: x.str.replace( r'[^\w\s]', '').str.lower()).replace(['rating', ''],
np.nan, regex=True)
df1= df1.dropna().reset_index(drop=True)

names = df1['Review'].tolist()

a = []
for sentence in names:
    score = analyser.polarity_scores(sentence)
    a.append(score)

dataFrame=pd.DataFrame(a)
c= dataFrame.drop(columns=['neg', 'neu', 'pos'])

d = pd.concat([df1, c], axis = 1)

#untuk ulasan positif
d['score']= d['compound'].apply(lambda x: 1 if x >= 0.05 else 0)
g = d[d['lol'].isin([0]) == False]
g.to_excel('Ulasan Positif.xlsx')

#untuk ulasan negatif
d['score']= d['compound'].apply(lambda x: 1 if x <= -0.05 else 0)
g = d[d['lol'].isin([0]) == False]
g.to_excel('Ulasan Negatif.xlsx')
```

Lampiran 3. *Script* Klasifikasi dengan Menggunakan *Naïve Bayes Classifier*

```
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
gnb = MultinomialNB()
cvec = CountVectorizer()

# splitting dataframe by row index
df_1 = d.iloc[:25000,:]
df_2 = d.iloc[25001:,:]

ccd = df_1['Review'].tolist()

txt_vec = cvec.fit_transform(ccd).toarray()
txt_vec.shape

label = df_1['score']
features = txt_vec
Jika:
# Perbandingan data latih sebesar 60% dan data uji 40% :
X_train, X_test, y_train, y_test = train_test_split(features, label, test_size=0.40,
random_state=42)

# Perbandingan data latih sebesar 70% dan data uji 30% :
X_train, X_test, y_train, y_test = train_test_split(features, label, test_size=0.30,
random_state=42)

# Perbandingan data latih sebesar 80% dan data uji 20% :
X_train, X_test, y_train, y_test = train_test_split(features, label, test_size=0.20,
random_state=42)

# Perbandingan data latih sebesar 85% dan data uji 15% :
X_train, X_test, y_train, y_test = train_test_split(features, label, test_size=0.15,
random_state=42)

gnb.fit(X_train, y_train)
y_pred1 = gnb.predict(X_test)
y_pred1
print("=== Classification Report ===")
print(classification_report(y_test, y_pred1, digits=4))
print(confusion_matrix(y_test, y_pred1))
```

Lampiran 4. *Script Association Rule* (Ulasan Positif)

```
#import library yang dibutuhkan
import pandas as pd
import numpy as np
import nltk
import re
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize

df = pd.read_excel('Ulasan Positif.xlsx')

df=df.dropna().reset_index(drop=True)
df = df.replace('connecting','connect', regex=True).replace('connection','connect',
      regex=True).replace('classes','class', regex=True).replace('problems','problem',
      regex=True).replace('meetings','meeting', regex=True).replace('working','work',
      regex=True).replace('audios','audio', regex=True).replace('videos','video',
      regex=True).replace('using','use', regex=True)

def tokenize(txt):
   tokens=re.split('\W+', txt)
   return tokens
df['Tokenized'] = df['Review'].apply(lambda x: tokenize(x))

stopwords_english = stopwords.words('english')
stop_words = ['good', 'easy', 'also', 'great', 'best', 'use', 'excellent', 'like', 'helpful',
              'love', 'please', 'much', 'app', 'zoom', 'ok', 'amazing', 'one', 'thanks',
              'really', 'get', 'nice', 'many', 'always', 'thank', 'bad', 'useful',
              'super', 'awesome', ',', '.', '', 'properly']+ list(stopwords_english)

def remove_stopwords(txt_tokenized):
   txt_clean = [word for word in txt_tokenized if word not in stop_words]
   return txt_clean

df['Stopword'] = df['Tokenized'].apply(lambda x: remove_stopwords(x))
df = df['Stopword']

from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules

te = TransactionEncoder()
te_ary = te.fit(df).transform(df)
df1 = pd.DataFrame(te_ary, columns = te.columns_)
```

```
frequent_itemsets = apriori(df1, min_support = 0.002, use_colnames=True)
frequent_itemsets

res = association_rules(frequent_itemsets, metric='lift', min_threshold=1)
res

res1=res[['antecedents', 'consequents', 'support', 'confidence', 'lift']]
res1

res1.to_excel('AR Positif.xlsx')
```

Lampiran 5. *Script Association Rule* (Ulasan Negatif

```
#import library yang dibutuhkan
import pandas as pd
import numpy as np
import nltk
import re
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize

df = pd.read_excel('Ulasan Negatif.xlsx')

df=df.dropna().reset_index(drop=True)
df = df.replace('connecting','connect', regex=True).replace('connection','connect',
      regex=True).replace('classes','class', regex=True).replace('problems','problem',
      regex=True).replace('meetings','meeting', regex=True).replace('working','work',
      regex=True).replace('audios','audio', regex=True).replace('videos','video',
      regex=True).replace('using','use', regex=True)

def tokenize(txt):
   tokens=re.split('\W+', txt)
   return tokens
df['Tokenized'] = df['Review'].apply(lambda x: tokenize(x))

stopwords_english = stopwords.words('english')
         stop_words = ['good', 'worst', 'bad', 'please', 'use', 'even', 'like',
                       'ever', 'also', 'much', 'app', 'zoom', 't', 'hai', '1',
                       'star', 'get', 'nice', 'many', 'always', 'experience',
                       'hate', 'every', ',', '.', '', 'properly']+ list(stopwords_english)

def remove_stopwords(txt_tokenized):
   txt_clean = [word for word in txt_tokenized if word not in stop_words]
   return txt_clean

df['Stopword'] = df['Tokenized'].apply(lambda x: remove_stopwords(x))
df = df['Stopword']

from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules

te = TransactionEncoder()
te_ary = te.fit(df).transform(df)
df1 = pd.DataFrame(te_ary, columns = te.columns_)
```

```
frequent_itemsets = apriori(df1, min_support = 0.004, use_colnames=True)
frequent_itemsets

res = association_rules(frequent_itemsets, metric='lift', min_threshold=1)
res

res1=res[['antecedents', 'consequents', 'support', 'confidence', 'lift']]
res1

res1.to_excel('AR Negatif.xlsx')
```