

**ANALISIS REVIEW PENGGUNA GOOGLE MEET DAN
ZOOM CLOUD MEETING MENGGUNAKAN ALGORITMA
NAÏVE BAYES DAN K-NEAREST NEIGHBOR**



TESIS

MUHAMMAD REZKI
14002242

PROGRAM STUDI ILMU KOMPUTER (S2)
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
NUSA MANDIRI
JAKARTA
2020

**ANALISIS REVIEW PENGGUNA GOOGLE MEET DAN
ZOOM CLOUD MEETING MENGGUNAKAN ALGORITMA
NAÏVE BAYES DAN K-NEAREST NEIGHBOR**



TESIS

Diajukan Sebagai salah satu syarat untuk memperoleh gelar
Magiester Ilmu Komputer (M.Kom)

MUHAMMAD REZKI

14002242

**PROGRAM STUDI ILMU KOMPUTER (S2)
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
NUSA MANDIRI
JAKARTA
2020**

SURAT PERNYATAAN ORISINILITAS

Yang bertanda tangan di bawah ini:

Nama : Muhammad Rezki
NIM : 14002242
Program Studi : Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Data Mining

Dengan ini menyatakan bahwa Tesis yang telah saya buat dengan judul: "Analisis Review Pengguna Google Meet Dan Zoom Cloud Meeting Menggunakan Algoritma *Naive Bayes* Dan *K-Nearest Neighbor*", adalah asli (orisinal) atau tidak plagiat (menjiplak) dan belum pernah diterbitkan/dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya tanpa ada paksaan dari pihak manapun juga. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa Tesis yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri dicabut/dibatalkan.

Jakarta, 3 Agustus 2020

Yang menyatakan,

Muhammad Rezki



PERSETUJUAN DAN PENGESAHAN TESIS

Tesis ini diajukan oleh :

Nama : Muhammad Rezki

NIM : 14002242

Program Studi : Ilmu Komputer

Jenjang : Strata Dua (S2)

Konsentrasi : Data Mining

Judul Tesis : “Analisis Review Pengguna Google Meet Dan Zoom Cloud Meeting Menggunakan Algoritma *Naïve Bayes* Dan *K-Nearest Neighbor*”

Telah dipertahankan pada periode 2020-1 dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Ilmu Komputer (M.Kom) pada Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri)

Jakarta, 04 Agustus 2020

PEMBIMBING TESIS

Pembimbing I : Dr. Lindung Parningotan Manik, M.T.I. 

Pembimbing II : Dr. Dwiza Riana, S.Si, MM, M.Kom 

DEWAN PENGUJI

Penguji I : Dr. Didi Rosiyadi, M.Kom 

Penguji II : Dr. Rifki Sadikin, M.Kom 

Penguji III/
Pembimbing : Dr. Lindung Parningotan Manik, M.T.I. 

KATA PENGANTAR

Puji syukur alhamdulillah, penulis panjatkan kehadiran Allah, SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan Tesis ini dengan baik. Dimana Tesis ini penulis sajikan dalam bentuk buku yang sederhana. Adapun judul Tesis, yang penulis ambil sebagai berikut “Analisis Review Pengguna Google Meet Dan Zoom Cloud Meeting Menggunakan Algoritma *Naïve Bayes* Dan *K-Nearest Neighbor*”.

Tujuan penulisan tesis ini dibuat sebagai salah satu untuk mendapatkan gelar Ilmu Komputer (S2) (M.Kom) pada Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri.

Tesis ini diambil berdasarkan hasil penelitian atau riset mengenai analisis review pengguna Google Meet Dan Zoom Cloud Meeting yang penulis lakukan pada google play, Penulis juga lakukan mencari dan menganalisis berbagai macam sumber referensi, baik dalam bentuk jurnal ilmiah, buku-buku literatur, internet, dll yang terkait dengan pembahasan pada tesis ini.

Penulis menyadari bahwa tanpa bimbingan dan dukungan dari semua pihak dalam pembuatan tesis ini, maka penulis tidak dapat menyelesaikan tesis ini tepat pada waktunya. Untuk itu ijinkanlah penulis kesempatan ini untuk mengucapkan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT yang selalu mencurahkan nikmat dan rahmat-Nya kepada penulis sehingga penulis dapat menyelesaikan tesis ini tepat pada waktunya
2. Bapak Dr. Lindung Parningotan Manik, M.T.I serta Ibu Dr. Dwiza Riana, S.Si, MM, M. Kom selaku pembimbing tesis yang telah menyediakan waktu, pikiran dan tenaga dalam membimbing penulis dalam menyelesaikan tesis ini.
3. Orang tua yang senantiasa tanpa henti selalu mendoakan dan selalu tiada henti memberikan motivasi, nasehat dan segala perhatiannya.
4. Seluruh staf pengajar (dosen) Program Studi Ilmu Komputer STMIK Nusa Mandiri yang telah memberikan pelajaran yang berarti bagi penulis selama menempuh studi.
5. Seluruh staf dan karyawan Program Studi Ilmu Komputer STMIK Nusa Mandiri yang telah melayani penulis dengan baik selama kuliah.

6. Teman-teman STMIK Nusa Mandiri angkatan tahun 2018 kelas 14.4B.01, yang selalu kompak dalam memberikan dukungan dan semangat untuk menyelesaikan tesis ini.

Serta semua pihak yang terlalu banyak untuk penulis sebutkan satu persatu sehingga terwujudnya penulisan tesis ini. Penulis menyadari bahwa penulisan tesis ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan karya ilmiah yang penulis hasilkan untuk yang akan datang.

Akhir kata semoga tesis ini dapat bermanfaat bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 3 agustus 2020



Muhammad Rezki

Penulis

**SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH
UNTUK KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Muhammad Rezki
NIM : 14002242
Program Studi : Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Data Mining
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, dengan ini menyetujui untuk memberikan ijin kepada pihak Program Studi Ilmu Komputer (S2) Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri) Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah kami yang berjudul : “Analisis Review Pengguna Google Meet Dan Zoom Cloud Meeting Menggunakan Algoritma *Naive Bayes* Dan *K-Nearest Neighbor*” beserta perangkat yang diperlukan (apabila ada).

Dengan Hak Bebas Royalti Non-Eksklusif ini pihak STMIK Nusa Mandiri berhak menyimpan, mengalih-media atau bentuk-kan, mengelolanya dalam pangkalan data (*database*), mendistribusikannya dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak STMIK Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya

Jakarta, 3 Agustus 2020

Yang menyatakan,

 
Muhammad Rezki

ABSTRAK

Nama : Muhammad Rezki
NIM : 14002242
Program Studi : Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Data Mining
Judul : Analisis Review Pengguna Google Meet Dan Zoom Cloud Meeting Menggunakan Algoritma Naïve Bayes Dan K-Nearest Neighbor

Penilaian Sebuah Aplikasi di *Playstore* memiliki tujuan untuk memberikan ulasan tentang kelebihan dan kekurangan dalam penggunaan aplikasi khususnya *virtual video conference*. Untuk mengetahui sejauh mana analisa review pengguna aplikasi Google Meet dan Zoom Cloud Meeting berdasarkan pemberian jumlah bintang dengan menggunakan teknik klasifikasi yaitu perbandingan Algoritma *Naïve Bayes* dan *k-Nearest Neighbor* (k-NN) dengan *feature* optimasi *SMOTE Upsampling* dan PSO. Hasil pengujian dengan metode 5 kelas mendapatkan hasil akurasi sebesar 77,92 % dengan menggunakan metode algoritma *K-Nearest Neighbor* yang ditambah dengan *Feature Smote* dan PSO untuk pengguna Zoom Cloud Meeting, sedangkan untuk pengguna Google Meet dengan metode yang sama yaitu *K-Nearest Neighbor* yang ditambah dengan *Feature Smote* dan PSO hanya mampu mendapat tingkat akurasi sebesar 73,84 %. sedangkan untuk pengujian dengan metode 2 kelas mendapatkan nilai akurasi tertinggi sebesar 91,56 % untuk pengguna Zoom Cloud Meeting dengan menggunakan algoritma *K-Nearest Neighbor* (k-NN).

Kata Kunci:

Text Mining, Algoritma Naive Bayes, k-Nearest Neighbor (k-NN), SMOTE Upsampling dan Particle Swarm Optimization (PSO)

ABSTRACT

Nama : Muhammad Rezki
NIM : 14002242
Program Studi : Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Data Mining
Judul : *Analysis of Google Meet User Review and Zoom Cloud Meeting Using the Naïve Bayes Algorithm and K-Nearest Neighbor*

Assessment of an application on Playstore aims to provide an overview of the advantages and disadvantages of using applications, especially virtual video conferencing. To find out the extent to which the analysis of user reviews for the Google Meet and Zoom Cloud Meeting applications is based on giving the number of stars using a classification technique, namely the comparison of the Naïve Bayes Algorithm and the k-Nearest Neighbor (k-NN) with the SMOTE Upsampling and PSO optimization feature. The test results with the 5 class method get an accuracy of 77.92% using the K-Nearest Neighbor algorithm method plus the Smote and PSO feature for Zoom Cloud Meeting users, while for Google Meet users with the same method, namely the K-Nearest Neighbor. coupled with the Smote Feature and PSO are only able to get an accuracy rate of 73.84%. while for testing with the 2 class method, the highest accuracy value is 91.56% for Zoom Cloud Meeting users using the K-Nearest Neighbor (k-NN) algorithm.

Keywords:

Text Mining, Algoritma Naive Bayes, k-Nearest Neighbor (k-NN), SMOTE Upsampling dan Particle Swarm Optimization (PSO)

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
LEMBAR PERNYATAAN ORISINALITAS	iii
HALAMAN PERSETUJUAN DAN PENGESAHAN TESIS	iv
KATA PENGANTAR	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	vi
ABSTRAK	vii
ABSTRACK	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1. Latar Belakang Penelitian	1
1.2. Identifikasi Masalah	4
1.3. Tujuan Penelitian	4
1.4. Ruang Lingkup Penelitian	5
1.5. Hipotesa	6
1.6. Sistematika Penulisan	6
BAB II LANDASAN TEORI.....	7
2.1. Tinjauan Pustaka.....	7
2.2. Tinjauan Studi.....	20
2.3. Tinjauan Organisasi/Objek Penelitian	25
BAB III METODOLOGI	27
3.1 Metode Pengumpulan Data.....	28
3.2. Teknik Analisa dan Pengujian Data.....	29
3.3 Instrumen Penelitian	30
3.4. Model Yang diusulkan.....	63
3.5 Model Yang diusulkan.....	31
3.5. Evaluasi dan Validasi Hasil	32
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	33
4.1. Pengumpulan Data	33
4.2. Pengolahan Data	34
4.3. Metode Yang Diusulkan	37
4.4. Evaluasi Model Klasifikasi.....	38
4.5. Evaluasi Dan Hasil Pengujian.....	43
BAB V PENUTUP.....	59
5.1. Kesimpulan	59
5.2. Saran	60
DAFTAR PUSTAKA	61
DAFTAR RIWAYAT HIDUP.....	65
LEMBAR KONSULTASI BIMBINGAN	66
LAMPIRAN.....	67

DAFTAR TABEL

Tabel 1.1. Penilaian rating bintang pengguna aplikasi	1
Tabel 2.1. Tahap Tokenizing	11
Tabel 2.2. Tahap Filtering.....	12
Tabel 2.3. Teknik text mining pre-processing Gata Framework.....	17
Tabel 2.4. Penelitian Terkait.....	24
Tabel 3.1. Review Pengguna Google Meet dan Zoom Cloud Meeting.....	30
Tabel 4.1. Jumlah Data Review Pengguna Bintang 1 – 5.....	34
Tabel 4.2. Jumlah Data Review Pengguna Bintang 1 dan 5.....	34
Tabel 4.3. Grafik Jumlah Data Review Aplikasi Google Meet dan Zoom Cloud Meeting.....	35
Tabel 4.4. Perbandingan Nilai Accuracy dan Kappa dengan NB dan k-NN (metode pertama: 5 Kelas, Bintang 1 – Bintang 5).....	40
Tabel 4.5. Hasil Pengujian Algoritma Naïve Bayes dan k-Nearest Neighbor dengan feature SMOTE Upsampling dan PSO (metode pertama: 5 Kelas, Bintang 1- Bintang 5).....	43
Tabel 4.6. Perbandingan Nilai Accuracy dan AUC dengan Algoritma NB dan k-NN (metode kedua: 2 Kelas, Bintang 1 dan Bintang 5).....	43
Tabel 4.7. Perbandingan Nilai Accuracy dan AUC dengan Algoritma NB dan k-NN dengan feature SMOTE Upsampling dan PSO (metode kedua: 2 Kelas, Bintang 1 dan Bintang 5)	43
Tabel 4.8. Hasil review pengguna Google Meet dengan Algoritma Naïve Bayes.....	44
Tabel 4.9. Hasil review pengguna Google Meet dengan Algoritma Naïve.....	45
Tabel 4.10. Hasil review pengguna Zoom Cloud Meeting dengan Algoritma Naïve Bayes.....	46
Tabel 4.11. Hasil review pengguna Zoom Cloud Meeting dengan Algoritma k-Nearest Neighbor (k-NN).....	46
Tabel 4.12. Hasil review pengguna Google Meet dengan Algoritma Naïve Bayes+SMOTE+PSO	47
Tabel 4.13. Hasil review pengguna Google Meet dengan Algoritma k-Nearest Neighbor (k-NN)+SMOTE+PSO.....	48
Tabel 4.14. Hasil review pengguna Zoom Cloud Meeting dengan Algoritma Naïve Bayes+SMOTE+PSO.....	49

Tabel 4.15. Hasil review pengguna Zoom Cloud Meeting dengan Algoritma k-Nearest Neighbor (k-NN).....	50
Tabel 4.16. Hasil Pengujian Accuracy dan Kappa.....	52
Tabel 4.17. Hasil Review Pengguna Google Meet dengan algoritma Naïve Bayes.....	53
Tabel 4.18. Hasil Review Pengguna Google Meet dengan algoritma K-Nearest Neighbor.....	54
Tabel 4.19. Hasil Review Pengguna Google Meet dengan algoritma Naïve Bayes feature selection SMOTE dan feature optimasi PSO.....	54
Tabel 4.20. Hasil Review Pengguna Google Meet dengan algoritma K-Nearest Neighbor (k-NN) dengan feature selection SMOTE dan feature optimasi PSO	54
Tabel 4.21. Hasil Review Pengguna Zoom Cloud Meeting dengan algoritma Naive Bayes.....	55
Tabel 4.22. Hasil Review Pengguna Zoom Cloud Meeting dengan algoritma k-Nearest Neighbor (k-NN).....	55
Tabel 4.23. Hasil Review Pengguna Zoom Cloud Meeting dengan algoritma Naïve Bayes dengan feature selection SMOTE dan feature optimasi PSO	56
Tabel 4.24. Hasil Review Pengguna Zoom Cloud Meeting dengan algoritma k-Nearest Neighbor (k-NN) dengan feature selection SMOTE dan feature optimasi PSO	56
Tabel 4.25. Perbandingan nilai Akurasi dan Kappa.....	57
Tabel 4.26. Perbandingan nilai Akurasi dan AUC.....	57
Tabel 4.27. Peningkatan Accuracy sebelum dan sesudah menggunakan SMOTE dan PSO (Metode Pertama – 5 Kelas).....	58
Tabel 4.28. Peningkatan Accuracy sebelum dan sesudah menggunakan SMOTE dan PSO (Metode Kedua – 2 Kelas).....	59

DAFTAR GAMBAR

Gambar 2.1 Proses Knowledge Discovery in Database (KDD).....	14
Gambar 2.2 Confusion matrix.....	15
Gambar 2.3. Kerangka Pemikiran.....	20
Gambar 3.1. Tahapan Penelitian.....	29
Gambar 3.2. Contoh review komentar pengguna Zoom Cloud meeting.....	30
Gambar 4.1. Tampilan tools Gata Framework Textmining.....	36
Gambar 4.2. Hasil pre-processing dengan menggunakan tools Gataframework untuk pengguna Google Meet.....	37
Gambar 4.3. Hasil pre-processing dengan menggunakan tools Gataframework untuk pengguna Zoom Cloud Meeting.....	37
Gambar 4.4. Dataset Google Meet yang akan digunakan dalam RapidMiner.....	38
Gambar 4.5. Dataset Zoom Cloud Meeting yang akan digunakan dalam RapidmMiner.....	38
Gambar 4.6. Desain Model menggunakan Algoritma NB dan K-NN.....	39
Gambar 4.7. Desain Model Algoritma NB dan k-NN dengan feature SMOTE dan PSO.....	41
Gambar 4.8. Desain Evaluasi 10 Fold Cross-Validation untuk algoritma Naïve Bayes.....	41
Gambar 4.9. Desain Evaluasi 10 Fold Cross-Validation untuk algoritma K-Nearest Neighbor.....	41
Gambar 4.10. Desain Evaluasi Process Documents.....	42
Gambar 4.11. Desain model PSO Algoritma Naïve Bayes.....	42

DAFTAR LAMPIRAN

Daftar Riwayat Hidup	65
Lembar Konsultasi Bimbingan	66
Lampiran 1. Dataset Google Meet	68
Lampiran 2. Dataset Zoom Cloud Meeting	91

BAB I

PENDAHULUAN

1.1. Latar Belakang Penulisan

Saat ini seluruh dunia sedang menghadapi wabah penyakit menular yaitu virus Covid 19. Pembatasan sosial atau menjaga jarak adalah serangkaian tindakan pengendalian infeksi nonfarmasi yang dimaksudkan untuk menghentikan atau memperlambat penyebaran penyakit menular tersebut. Sehingga seluruh masyarakat diharapkan untuk beraktifitas dirumah untuk menghentikan penyebaran virus Covid 19. Agar tetap bisa menjalankan aktifitas dirumah diperlukan *virtual meet* untuk berkomunikasi sesama team atau rekan kerja. Saat ini *virtual meet* telah banyak dipakai.

Google Play adalah layanan konten digital milik Google yang terdiri dari toko produk-produk online seperti musik/lagu, buku, aplikasi, permainan, ataupun pemutar media. Google Play dapat diakses melalui web, aplikasi android (PlayStore), dan Google TV. Dalam Google Play, selain *virtual meet* juga dilengkapi adanya fitur penilaian bagi pelanggan untuk memberikan ulasan tentang kelebihan dan kekurangan dalam penggunaan aplikasi *virtual video conference*, ulasan pengguna bisa bermacam-macam bentuknya, mulai dari kalimat yang halus maupun kasar juga ada tergantung penilaian masing-masing penggunanya. Pengguna google play, selain memberikan ulasan juga bisa memberikan penilaian berupa skor bintang (antara 1 – 5) tergantung penilaian yang diberikan oleh pengguna tersebut dan terakhir pengguna juga bisa memberikan skor “like” terhadap komentar seseorang yang memang bisa mewakili ketertarikan atau ketidaksukaan seseorang terhadap aplikasi *virtual meet* yang digunakan.

Tabel 1.1 Penilaian rating bintang pengguna aplikasi

No	Bintang	keterangan
1	★ ★ ★ ★ ★	Sangat bagus
2	★ ★ ★ ★	Cukup
3	★ ★ ★	Bagus
4	★ ★	Kurang bagus
5	★	Sangat kurang

Sumber: Telah diolah kembali berdasarkan pencarian di google play

Ulasan dari pengguna sering digunakan sebagai alat yang efektif dan efisien dalam menemukan informasi terhadap suatu produk atau jasa. Bahwa kebanyakan dari pengguna internet bergantung pada rekomendasi *word-of-mouth* (opini) sebelum menggunakan suatu produk, karena review dari pengguna lain dapat menyediakan informasi terbaru dari produk tersebut berdasarkan perspektif pengguna-pengguna lain yang sudah menggunakan produk tersebut. Namun untuk memantau dan mengorganisir opini dari masyarakat bukanlah hal yang mudah. Opini yang dimuat di media sosial jumlahnya terlalu banyak untuk diproses secara manual. Oleh sebab itu peneliti ingin mengetahui sejauh mana analisis review pengguna aplikasi *virtual meet* berdasarkan pemberian jumlah bintang (*) setelah menggunakan aplikasi tersebut.

Untuk penelitian ini, fokus utamanya adalah review pengguna aplikasi Zoom Cloud Meeting dan Google Meet yang ada di situs Google Play dimana untuk pengambilan datanya pada Bulan Maret 2020 – Mei 2020 (kurun waktu 3 bulan). Citra merek yang baik tentunya akan membentuk opini yang baik pula dari konsumen tentang suatu produk/jasa. Review pengguna Zoom Cloud Meeting maupun Google Meet bisa jadi dipengaruhi oleh beberapa hal yang belum menjadi perhatian baik dari pihak Zoom Cloud Meeting maupun Google Meet. Hal ini mungkin terjadi karena adanya beberapa faktor yang harus diperbaiki dan belum diketahui oleh pihak Zoom meeting maupun Google Meet, agar kualitas dari dari aplikasi semakin baik.

Salah satu analisis *text mining* yaitu analisis sentimen dapat diaplikasikan pada perusahaan yang mengeluarkan produk dan jasa yang menyediakan layanan untuk menerima pendapat (*feedback*) dari konsumen untuk produk tersebut. Analisis sentimen diaplikasikan untuk mengelompokkan *feedback* positif, negatif, dan netral dari konsumen sehingga membantu tugas perusahaan untuk meninjau kembali kekurangan produk mereka. Apabila ditemukan adanya sentiment negatif, maka perusahaan dapat dengan cepat mengambil tindakan untuk memperbaiki produknya. Analisis sentimen merupakan riset komputasional dari opini, sentimen, dan emosi yang diekspresikan secara tekstual [1].

Terdapat beberapa penelitian yang sudah dilakukan dalam melakukan klasifikasi sentimen terhadap review yang tersedia secara online diantaranya, Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan *Naïve Bayes* dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile) [1]. Analisis Sentimen Data Komentar Sosial Media Facebook Dengan K-Nearest Neighbor (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekspres Indonesia)[2]. Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan *Metode Support Vector Machine* dan *Lexicon Based Features* [3]. Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma *Naïve Bayes* [4]. Analisa Model *Support Vector Machine Text mining* Pada Komentar Positif Dan Negatif Untuk Review Perbandingan Whatsapp Vs Bbm [5]. Aplikasi Web Untuk Analisis Sentimen Pada Opini Produk Dengan Metode *Naïve Bayes Classifier* [6]. Dari beberapa penelitian terkait, kami mengambil metode *Naïve Bayes* dan *K-Nearest Neighbor* berbasis *feature selection*, sehingga penelitian ini bertujuan untuk membandingkan 2 metode algoritma dan membandingkan 2 aplikasi *virtual meet* yaitu Google Meet dan Zoom Cloud Meeting dengan sumber data *review* pengguna pada google play berdasarkan *rating* bintang yang didapatkan.

Dari serangkaian latar belakang yang telah diuraikan diatas, peneliti akan melakukan analisis review pengguna Google Meet dan Zoom Cloud Meeting berdasarkan perolehan bintang yang didapatkan dari setiap komentar dengan menggunakan metode *Naïve Bayes* dan *K-Nearest Neighbor* sebagai pembandingnya dan penggunaan *feature selection* dan *feature optimasi* untuk ujicoba penelitiannya. Dari hasil penelitian ini, tentunya dapat diketahui aplikasi virtual meet manakah yang memiliki tingkat akurasi paling tinggi dari skor perolehan penilaian bintang yang sering diberikan pengguna sesuai dengan hasil penelitian menggunakan model algoritma *Naïve Bayes* dan *K-Nearest Neighbor*.

1.2. Identifikasi Masalah

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

1. Bagaimanakah gambaran umum data *review* pengguna Google Meet dan Zoom Cloud Meeting pada *Google Play* dengan pengambilan data dalam jangka waktu 3 bulan dengan menggunakan algoritma klasifikasi untuk teks berbahasa Indonesia berdasarkan jumlah skor bintang dengan menggunakan 2 kelas berbeda yaitu pertama untuk penilaian skor bintang 1 sampai dengan bintang 5, kedua penilaian skor bintang 1 dan bintang 5 saja?
2. Apakah ada perubahan signifikan hasil perhitungan antara Algoritma *Naïve Bayes* dan *K-Nearest Neighbor* jika menggunakan *feature selection Synthetic Minority Over-Sampling Technique (SMOTE)* dan *feature optimasi Particle Swarm Optimization (PSO)* atau tidak menggunakan *feature* tersebut? Kemudian Algoritma manakah yang memiliki nilai akurasi paling tinggi?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian yang dilakukan dalam tesis ini sebagai berikut:

1. Mengklasifikasikan analisis review pengguna 2 aplikasi *virtual meeting* yaitu Google Meet dan Zoom Cloud Meeting dengan proses *Text Mining* menggunakan metode Algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dengan menambahkan *feature selection Synthetic Minority Over-Sampling Technique (SMOTE)* dan *feature optimasi Particle Swarm Optimization (PSO)* sehingga bisa menghasilkan nilai akurasi tertinggi sebelum dan sesudah menggunakan *feature* tambahan tersebut.
2. Membantu pengembangan dan penerapan teori para peneliti lainnya yang berkaitan dengan analisis review pengguna Google Meet dan Zoom Cloud Meeting pada *google play* berdasarkan jumlah skor bintang yang diberikan pengguna dengan membandingkan hasil penilaian 5 kelas (bintang 1 sampai dengan bintang 5) dan 2 kelas (bintang 1 dan bintang 5 saja).

1.4 Ruang Lingkup Penelitian

Ruang lingkup masalah sudah ditentukan untuk menghindari perluasan pembahasan dalam penelitian ini, antara lain:

1. Data yang akan diklasifikasi hanya *review* pengguna Google Meet dan Zoom Cloud Meeting pada *Google Play* dalam kurun waktu Bulan Maret 2020 sampai dengan Mei 2020 (3 bulan).
2. *Review* pada aplikasi yang akan diklasifikasi adalah teks berbahasa Indonesia (meliputi bahasa formal maupun non formal).
3. Metode yang digunakan untuk analisis *review* menggunakan Algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dengan menambahkan *feature selection Synthetic Minority Over-Sampling Technique (SMOTE)* dan *feature optimasi Particle Swarm Optimization (PSO)*, sehingga mampu mengetahui algoritma manakah yang memiliki tingkat akurasi paling tinggi beserta hasil analisisnya berdasarkan hasil perhitungan kedua algoritma tersebut dengan membandingkan hasil penilaian 5 kelas (bintang 1 sampai dengan bintang 5) dan hasil penilaian 2 kelas (bintang 1 dan bintang 5 saja).

1.5 Hipotesis

Hipotesis adalah dugaan awal terhadap sebuah penelitian yang dilakukan. Adapun hipotesis terhadap penelitian yang dilakukan, yaitu:

1. Memberikan gambaran umum kepada pengguna *Zoom Cloud Meeting* dan *Google Meet* pada *Google Play* dengan pengambilan data dalam jangka waktu 3 bulan dengan menggunakan algoritma klasifikasi untuk teks berbahasa Indonesia berdasarkan jumlah skor bintang.
2. Ada perubahan signifikan hasil perhitungan antara Algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dengan menggunakan *feature selection Synthetic Minority Over-Sampling Technique (SMOTE)* dan *feature optimasi Particle Swarm Optimization (PSO)*.

1.6 Sistematika Penulisan

Sistematika penulisan tesis mengenai analisis *review* pengguna Zoom Cloud Meeting dan Google Meet pada *google play* dapat disusun sebagai berikut:

Bab I : PENDAHULUAN

Pada bab ini berisi uraian fakta-fakta berkaitan permasalahan yang akan diteliti beserta rencana penelitian untuk mendapatkan solusi atas permasalahan tersebut.

Bab II : LANDASAN TEORI

Pada bab ini akan dijelaskan tentang landasan-landasan teoritis yang digunakan untuk melaksanakan penelitian dan uraian sistematis dari penelitian-penelitian terkait.

Bab III : METODOLOGI PENELITIAN

Pada bab ini membahas metode pengumpulan data dari eksperimen menguji Algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dengan menambahkan *feature selection SMOTE Upsampling* dan *feature optimasi PSO*.

Bab IV : HASIL PENELITIAN DAN PEMBAHASAN

Pada bab ini berisi tentang pembahasan yang dilakukan tentang tata cara *text mining* yang dilakukan, penggunaan algoritma *text mining* yaitu *Naïve Bayes* dan *K-Nearest Neighbor* dengan *feature selection Synthetic Minority Over-Sampling Technique (SMOTE)* dan *feature optimasi Particle Swarm Optimization (PSO)*, langkah-langkah perhitungan dari metode yang digunakan beserta nilai *Accuracy*, *Kappa* dan *AUC* yang dihasilkan dari analisis review pengguna *Zoom Cloud Meeting* dan *Google Meet* .

Bab V : PENUTUP

Pada bab ini berisi tentang kesimpulan dari penjelasan pada bab-bab sebelumnya dan saran untuk penelitian selanjutnya.

BAB II

LANDASAN/KERANGKA PEMIKIRAN

2.1 Tinjauan pustaka

Pada penelitian ini, tinjauan pustaka yang digunakan merupakan teori – teori yang menjadi landasan dalam penelitian ini. Selain itu kajian pustaka juga didapat melalui referensi dari jurnal – jurnal penelitian.

2.1.1. Analisis Sentimen

Analisis sentimen merupakan salah satu cabang penelitian dari Text Mining yang melakukan proses klasifikasi pada dokumen teks [7]. Analisis sentimen merupakan sebuah metode yang digunakan untuk mengekstrak data opini, memahami serta mengolah tekstual data secara otomatis untuk melihat sentimen yang terkandung dalam sebuah opini [8]. Analisis sentimen atau opinion mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini [9]. Analisis sentimen adalah bidang interdisipliner, sebuah bidang dimana pendekatan pemecahan masalahnya dengan menggunakan tinjauan dari berbagai sudut pandang ilmu serumpun secara relevan dan terpadu [10].

Analisis sentimen dapat dibedakan berdasarkan sumber datanya, beberapa level yang sering digunakan dalam penelitian analisis sentimen adalah analisis sentimen pada level dokumen dan analisis sentimen pada level kalimat [11]. Berdasarkan level sumber datanya Analisis Sentimen terbagi menjadi dua kelompok besar yaitu:

1. Coarse-grained Sentiment Analysis

Pada Sentiment Analysis Coarse-grained, Analisis Sentimen yang dilakukan adalah pada level dokumen. Secara garis besar fokus utama dari Sentiment Analysis jenis ini adalah menganggap seluruh isi dokumen sebagai sebuah sentiment positif atau sentiment negatif [12].

2. Fined-grained Sentiment Analysis

Fined-grained Sentiment Analysis adalah Sentiment Analysis pada level kalimat. Fokus utama fined-grained Sentiment Analysis adalah menentukan sentimen pada setiap kalimat [12].

2.1.2. Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar [13]. Data mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika [14]. Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* (KDD) secara keseluruhan [15]. Data mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data warehouse, atau penyimpanan informasi lainnya. *Data mining* berkaitan dengan bidang ilmu-ilmu lain, seperti database sistem, *data warehousing*, statistik, *machine learning*, *information retrieval*, dan komputasi tingkat tinggi [16].

Karakteristik data mining adalah sebagai berikut [16]:

1. Data mining berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
2. Data mining biasa menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk membuat hasil lebih dipercaya.
3. Data mining berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

2.1.3. Pengelompokan Data Mining

Menurut Larose dalam bukunya yang berjudul “*Discovering Knowledge in Data: An Introduction to Data Mining*”, data mining dibagi menjadi beberapa kelompok berdasarkan tugas/pekerjaan yang dapat dilakukan, yaitu [13]:

a. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

b. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan baris data (*record*) lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

c. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

d. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

e. Pengklasteran (*Clusterring*)

Pengklasteran merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas obyek-obyek yang memiliki kemiripan. Klaster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan record dalam klaster yang lain. Berbeda dengan klasifikasi, pada pengklasteran tidak ada variabel target. Pengklasteran tidak melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target, akan tetapi, algoritma pengklasteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (*homogen*), yang mana kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal.

f. Asosiasi

Tugas asosiasi dalam data mining adalah untuk menemukan atribut yang muncul dalam satu waktu. Salah satu implementasi dari asosiasi adalah market basket analysis atau analisis keranjang belanja, sebagaimana yang akan dibahas dalam tesis ini.

2.1.4. Tahapan proses data mining

Data mining lebih tepat disebut sebagai penambangan pengetahuan dari data, langkah-langkah penting dalam proses penambangan pengetahuan dari data secara umum sebagai berikut [16]:

1. Pembersihan data (*data cleaning*), yaitu proses menghapus data pengganggu (noise) yang dikatakan tidak konsisten atau tidak diperlukan.
2. Integrasi data (*data integration*), yaitu menggabungkan berbagai sumber data.
3. Pemilihan data (*data selection*), yaitu memilih data yang dipilih sesuai kebutuhan analisis.
4. Transformasi data (*data transformation*), yaitu proses transformasi data ke dalam format untuk diproses dan siap ditambang.
5. Penggalian data (*data mining*), yaitu menerapkan metode kecerdasan untuk ekstraksi pola.
6. Evaluasi pola (*pattern evaluation*), yaitu mengidentifikasi pola-pola yang menarik yang merepresentasikan pengetahuan.
7. Penyajian pola (*knowledge presentation*), yaitu teknik untuk memvisualisasikan pola pengetahuan ke pengguna.

2.1.5. Text Mining

Text mining (menambang teks) merupakan analisis teks dimana sumber data biasanya di dapatkan dari dokumen, dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan, keterkaitan dan kelas antar dokumen [17]. *Text mining* merupakan salah satu teknik yang dapat digunakan untuk melakukan klasifikasi, dimana text mining merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar [18]. Dalam penerapan text mining, terdapat beberapa langkah yang perlu dilakukan antara lain [18]:

1. Tokenizing

Tokenizing merupakan proses penguraian *Deskripsi* yang semula berupa kalimat menjadi kata. Contoh proses tokenizing pada sebuah kalimat dapat dilihat pada Tabel II.1

Tabel 2.1 Tahap *Tokenizing*

<i>Text Input</i>	<i>Hasil Token</i>

<i>Go until jurong point, crazy..</i>	<i>go</i>
<i>Available only in bugis n</i>	<i>until</i>
<i>great world la e buffet... Cine</i>	<i>jurong</i>
<i>there got amore wat...</i>	<i>point</i>
	<i>crazy</i>
	<i>available</i>
	...
	<i>wat</i>

Berdasarkan tabel 2.1 semua kata yang menyusun kalimat pada kolom “Text Input” dipotong berdasarkan kata yang menyusunnya seperti terlihat di kolom “Hasil Token” pada Tabel II.1.

2. Filtering

Filtering adalah tahap mengambil kata penting dari hasil proses token. Bisa menggunakan algoritma stop list atau word list. Filtering dapat juga diartikan sebagai proses mengambil kata – kata penting dari hasil proses token atau penghapusan stopwords. Stopwords merupakan kosa kata yang bukan merupakan ciri (kata unik) dari suatu dokumen. Untuk contoh tahap filtering terlihat pada Tabel II.2. Kolom “Hasil Token” pada Tabel II.2 adalah kata – kata yang berasal dari proses tokenizing sedangkan kata yang berada pada kolom “Hasil Filtering” adalah hasil setelah proses filtering yaitu menghilangkan kata yang tidak penting seperti kata penghubung “go”, “in”, dan “there”.

Tabel 2.2 Tahap Filtering

Hasil Token	Hasil Filtering
Go	Until
Until	Jurong
Jurong	Point
Point	Crazy
Crazy	Available
Available	Bugis
In	
Bugis	
....

wat	wat
-----	-----

3. Stemming

Stemming merupakan tahap untuk mencari root kata dari hasil filtering. Stemming adalah proses pemetaan dan penguraian berbagai bentuk (variants) dari suatu kata menjadi bentuk kata dasarnya (stem).

4. Tagging

Tagging merupakan tahap untuk mencari bentuk awal/root dari tiap kata lampau atau hasil dari proses stemming.

5. Tahap Analysing

Analyzing merupakan tahap penentuan seberapa jauh keterhubungan antar suatu kata atau term terhadap suatu dokumen atau kalimat dengan menghitung nilai/bobot keterhubungan. Algoritma TF/IDF digunakan dalam proses perhitungan bobot (W) terminologi kata. Algoritma ini digunakan untuk menghitung bobot setiap kata yang paling umum digunakan pada information retrieval.

2.1.6. Teknik klasifikasi

Teknik klasifikasi merupakan proses menemukan model (fungsi) yang mampu menjelaskan dan membedakan kelas-kelas atau konsep, dengan tujuan agar model yang diperoleh dapat digunakan untuk memprediksikan kelas atau objek yang memiliki label kelas tidak diketahui [19]. Tiap teknik klasifikasi menggunakan suatu algoritma pembelajaran untuk mendapatkan suatu model yang paling memenuhi hubungan antara himpunan atribut dan label kelas dalam data masukan. Berikut ini algoritma klasifikasi yang akan digunakan dalam penelitian ini antara lain:

1. *Naive Bayes*

Algoritma *Naive Bayes* adalah pengelompokan probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menghitung frekuensi dan kombinasi nilai-nilai dalam suatu kumpulan data yang diberikan [20]. *Naive Bayes* adalah metode data mining yang sederhana dalam pengolahan data yang besar. Dimana metode *Naive Bayes* mempunyai nilai akurasi yang tinggi jika diterapkan ke dalam database yang besar. *Naive Bayes* merupakan pendekatan statistik dalam pengenalan pola (pattern recognition) yang didasarkan pada asumsi penyederhanaan nilai atribut [21]. *Naive Bayes* didasarkan pada asumsi

penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu [22].

Perhitungan *Naïve Bayes* dapat dilakukan dengan langkah berikut ini [22]:

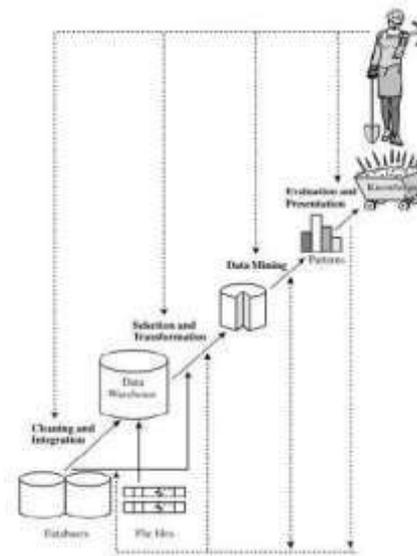
- a. Mencari nilai prioritas untuk tiap-tiap kelas dengan menghitung rata-rata tiap kelas dengan menggunakan persamaan.
 - b. Mencari nilai *Likelihood* untuk tiap-tiap kelas dengan menggunakan persamaan.
 - c. Mencari nilai posterior dari tiap kelas yang ada dengan menggunakan persamaan.
2. *K-Nearest Neighbor*

K-Nearest Neighbor merupakan metode untuk melakukan klasifikasi terhadap objek berdasarkan data training yang menggunakan jarak terdekat atau kemiripan terhadap objek tersebut. Pada fase pembelajaran, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi dari data pembelajaran [23]. Algoritma *K-Nearest Neighbor* merupakan sebuah algoritma yang sering digunakan untuk klasifikasi teks dan data. Metode KNN adalah metode yang melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan obyek tersebut. KNN termasuk algoritma *supervised learning*. *Supervised learning* merupakan suatu pembelajaran yang terawasi dimana jika output yang diharapkan telah diketahui sebelumnya [24]. *K-Nearest Neighbor* adalah salah satu metode pengenalan pola yang umum dan sering digunakan untuk proses klasifikasi sekelompok data karena tekniknya yang sederhana. Klasifikasi *K-Nearest Neighbor* mengkategorikan sebuah sampel data tidak berlabel dengan menggunakan label mayoritas dari sampel data tetangga terdekat (paling mirip) [25].

2.1.7. Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) merupakan proses analisa terstruktur untuk memperoleh informasi yang benar, baru, bermanfaat dan menemukan pola dari data yang besar dan kompleks. Data mining (DM) menjadi inti dari proses KDD, yakni dengan menggunakan algoritma tertentu untuk mengeksplorasi data, membangun model dan menemukan pola yang belum diketahui. Model digunakan untuk memahami fenomena data, analisa maupun

prediksi[26]. KDD merupakan kegiatan interdisiplin ilmu yang mengacu kepada metodologi untuk memperoleh pengetahuan dari data[27]. KDD merupakan suatu pola kegiatan non-trivial (tidak biasa) guna untuk mencari serta mengidentifikasi pola (pattern) yang terjadi pada sekumpulan data, dimana pola yang ditemukan bersifat sah (yang sebenarnya) baru dapat bermanfaat dan dapat dimengerti[28].



Gambar 2.1 Proses Knowledge Discovery in Database (KDD) [29]

2.1.8. Evaluasi dan Model Validasi

Confusion Matrix merupakan suatu metode yang digunakan untuk menghitung akurasi atau kinerja dari data mining. Terdapat 4 istilah dalam representasi hasil proses klasifikasi, yakni True Positive (TP) menyatakan observasi bernilai positif, dan diprediksi positif, False Negative (FN) menyatakan observasi bernilai positif, tetapi diprediksi negatif, True Negative (TN) menyatakan observasi bernilai negatif, dan diprediksi negatif, dan False Positive (FP) menyatakan observasi bernilai negatif, tetapi diprediksi positif[30].

		Predicted Values	
		Positive (0)	Negative (1)
Actual Values	Positive (0)	TP	FP
	Negative (1)	FN	TN

Gambar. 2.2 Confusion matrix [31]

Confusion Matrix menjadi sumber informasi apakah model yang digunakan memiliki performa yang baik atau tidak. Hal itu dapat dilihat pada angka - angka yang terdapat di dalamnya. Angka pada variabel TP (True Positif) dan variabel TN (True Negative) merepresentasikan total prediksi benar yang dilakukan oleh model. Sedangkan angka pada variabel FP (False Positive) dan variabel FN (False Negative) merepresentasikan total prediksi salah yang dihasilkan. Penghitungan performa model didapatkan dengan menghitung nilai accuracy, precision, recall, dan F1-Score, dengan rumus yang bisa dilihat pada persamaan.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$

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

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

$$F1 - Score = \frac{2x Precision x Recall}{Precision + Recall}$$

TP = Jumlah data kelas positif (0) diprediksi benar sebagai kelas positif (0)

FN = Jumlah data kelas positif (0) diprediksi salah sebagai kelas negatif (1)

TN = Jumlah data kelas negatif (1) diprediksi benar sebagai kelas negatif (1)

FP = Jumlah data kelas negatif (1) diprediksi salah sebagai kelas positif (0)

2.1.9. Gata Framework

Gataframework merupakan PHP Framework yang dibangun oleh Dr. Windu Gata, Mkom yang telah menjadi Dosen Tetap di STMIK Nusamandiri. Dengan pengalaman sebagai programmer sejak tahun 1995. Gataframework merupakan PHP Framework aplikasi berbasis web untuk pre-processing seperti penghilangan

kata-kata Indonesia, akronim Indonesia, bahasa gaul bahasa Indonesia dan lainnya yang dimaksudkan untuk membantu mahasiswa dan peneliti lain dalam melakukan penelitian di bidang text mining berbahasa Indonesia. Aplikasi ini dibangun menggunakan framework yang disebut Gata Framework (<http://www.gataframework.com>). Aplikasi ini merupakan alternatif dalam teks pre-processing Indonesia, aplikasi ini juga menyediakan fitur antarmuka program aplikasi (API) untuk mengirim data dari aplikasi eksternal. Sedangkan Gata Framework adalah kerangka kerja yang didasarkan pada bahasa pemrograman PHP yang dikembangkan dengan nama MTG Framework pada tahun 2012 dan mengubah namanya menjadi Gata Framework pada tahun 2017. Gata Framework telah mampu mengatasi berbagai masalah eksternal, yaitu usability, kapabilitas, respons, keamanan, keberadaan, dan keandalan, serta faktor internal, yaitu kemudahan sintaksis atau kode yang mudah digunakan dan telah menggunakan pola pemrograman Model View Controller (MVC). Saat ini, perangkat aplikasi dapat memproses pre-processing dalam bentuk satu formulir, atau mengunggah data dalam bentuk file Ms. Excell dengan template dan layanan web. Teknik yang biasa digunakan untuk pre-processing teks bahasa Indonesia adalah @notasi Penghapusan, Hapus URL, Tokenisasi: REGEXP, Transformasi Tidak (Negatif), Stemming Indonesia, dan Penghapusan Stopwords Indonesia. Penjelasan tentang setiap opsi teknik dalam penambangan teks Gata Framework dapat dilihat pada tabel dibawah ini.

Tabel 2.3. Teknik text mining pre-processing Gata Framework

No	Teknik	keterangan
1	@ Annotation Removal	Hapus tanda @ dan <i>Deskripsi</i> yang sering digunakan untuk menyapa atau menandai akun lain.
2	Transformation: Remove URL	Hapus URL dari tweet yang digunakan.
3	Tokenization: Regexp	Hapus tanda selain huruf dan juga hilangkan angka.
4	Transformation NOT	Hubungkan kata-kata yang memiliki makna

		terbalik seperti "tidak", "tidak", dll.
5	Indonesian Stemming	Mengembalikan kata menjadi kata dasar.
6	Indonesian Stopword removal	Hapus kata-kata yang tidak memiliki makna.

2.1.10. *Synthetic Minority Over-Sampling Technique (SMOTE)*

SMOTE merupakan salah satu metode *oversampling* untuk menangani masalah imbalance data dengan melakukan peningkatan pada data minor dengan cara duplikasi [32]. SMOTE menyeimbangkan kumpulan data dengan mengumpulkan data minoritas secara sintetis di ruang masukan berdasarkan informasi lingkungan mereka [33]. Data set pelatihan terdiri dari data point minoritas S_{min} dan data point mayoritas S_{maj} . Untuk setiap $(X_i, Y_i) \in S_{min}$, titik data terbanyak mengatur S_{maj} . Untuk masing-masing $(x_i, y_i) \in S_{min}$, SMOTE menghasilkan titik data minoritas baru di sepanjang segmen garis yang bergabung dengan x_i dan salah satu tetangga terdekat k yang dipilih secara acak.

$$X_{new} = X_i + (X_j - X_i) \times \delta$$

Dimana δ adalah bilangan acak antara $[0, 1]$. Label kelas x_{new} diatur ke $+1$ (y_{new}), label kelas minoritas. Titik data yang baru dihasilkan $\{(x_{new}, y_{new})\}$ kemudian ditambahkan ke set asli S dan digunakan untuk melatih classifier. Dari perumusan di atas jelas bahwa SMOTE menginformasikan oversampling menggeneralisasikan kelas minoritas dengan menciptakan daerah keputusan yang lebih besar dan kurang spesifik, daripada wilayah asli yang lebih kecil dan lebih spesifik.

2.1.11. *Particle Swarm Optimization (PSO)*

Particle Swarm Optimization (PSO) adalah metode optimasi global yang diperkenalkan oleh Kennedy dan Eberhart pada tahun 1995 berdasarkan penelitian terhadap perilaku kawanan burung dan ikan. Setiap partikel dalam Particle Swarm Optimization memiliki kecepatan partikel bergerak dalam ruang pencarian dengan kecepatan yang dinamis disesuaikan dengan perilaku historis mereka. Oleh karena itu, partikel memiliki kecenderungan untuk bergerak menuju daerah pencarian yang lebih baik selama proses pencarian [34]. Particle Swarm Optimization merupakan sebuah metode optimasi yang berdasarkan pada swarm intelligence atau yang

disebut algoritma behaviorally inspired yang dapat digunakan sebagai alternatif dari algoritma genetika [35]. Particle Swarm Optimization merupakan salah satu algoritma optimasi yang dapat digunakan untuk pengambilan keputusan. Algoritma ini bekerja dengan cara menghitung secara terus menerus calon solusi dengan menggunakan suatu acuan kualitas. Algoritma ini mengoptimasi permasalahan dengan cara menggerakkan partikel didalam ruangan permasalahan menggunakan fungsi tertentu [36]. Karena konsep sederhana, implementasi mudah, dan konvergensi cepat, particle swarm optimization (PSO) dapat diterapkan untuk berbagai aplikasi di berbagai bidang untuk memecahkan masalah optimasi. Implementasi PSO sangat mudah dan hanya dibutuhkan beberapa baris kode pemrograman, dan juga tidak dibutuhkannya operator matematika yang rumit, maka dari itu dapat mengefesienkan perhitungan baik dari segi memori yang di butuhkan dan juga dari segi kecepatan. PSO, sebagai alat optimasi, yang dapat membantu menentukan parameter optimum. Tetapi PSO memiliki ketergantungan yang sensitif pada parameter yang digunakan [36].

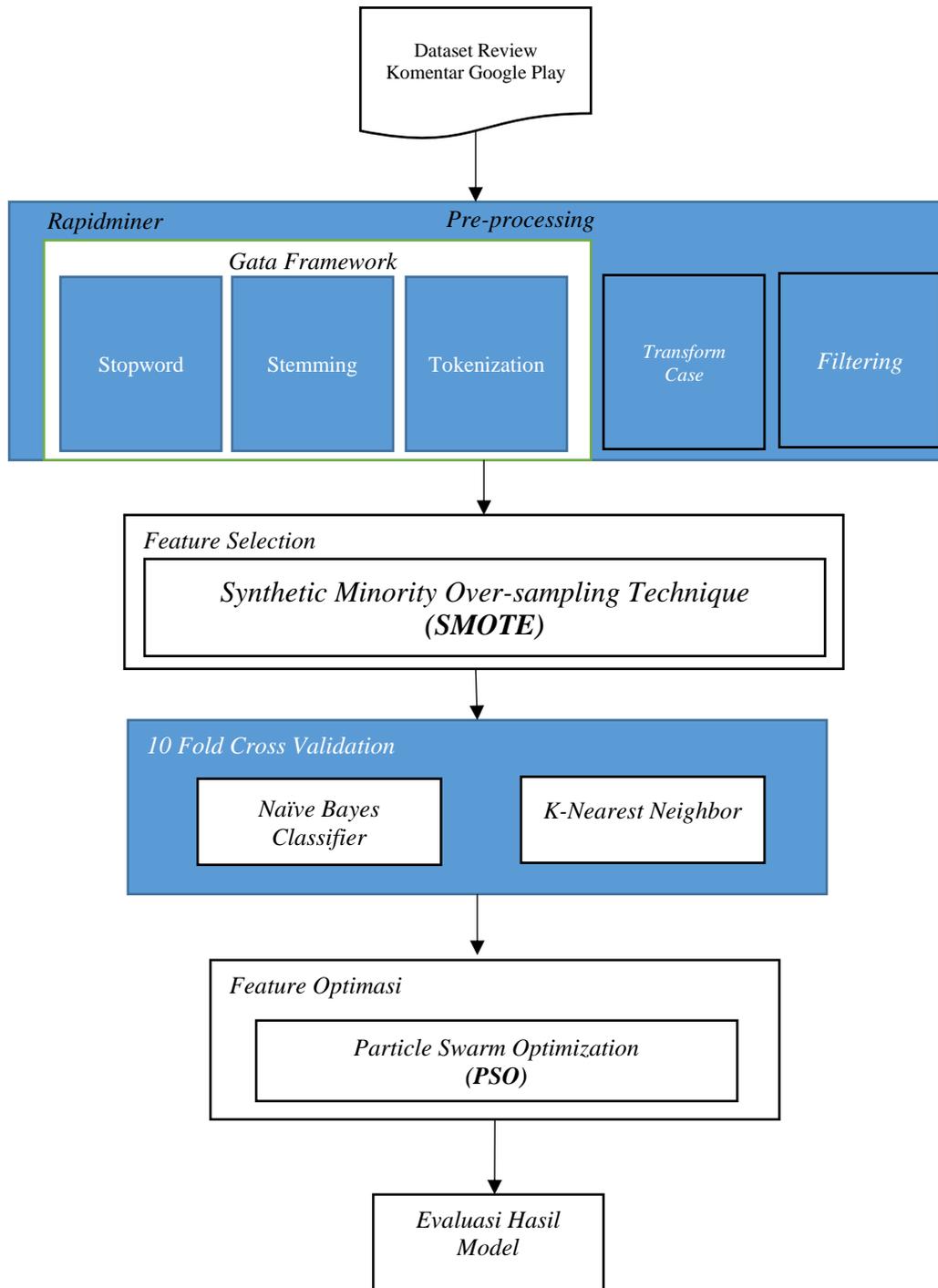
2.1.12. Rapidminer

Rapid Miner merupakan perangkat lunak yang dibuat oleh Dr. Markus Hofmann dari Institute of Technology Blanchardstown dan Ralf Klinkenberg dari rapid-i.com dengan tampilan GUI (*Graphical User Interface*) sehingga memudahkan pengguna dalam menggunakan perangkat lunak ini [29]. Perangkat lunak ini bersifat open source dan dibuat dengan menggunakan program Java di bawah lisensi GNU Public Licence dan Rapid Miner dapat dijalankan di sistem operasi manapun. RapidMiner merupakan software/perangkat lunak untuk pengolahan data. Dengan menggunakan prinsip dan algoritma data mining, RapidMiner mengekstrak pola-pola dari data set yang besar dengan mengkombinasikan metode statistika, kecerdasan buatan dan database [37].

2.1.13. Kerangka Pemikiran

Kerangka Pemikiran dalam penelitian ini menguraikan terhadap langkah penelitian yang akan dilakukan pada tesis ini sesuai pada gambar 2.5. Alur tahapan kerangka pemikiran penelitian ini, dapat dijelaskan sebagai berikut:

1. Pengumpulan data review analisis pengguna Google meet dan Zoom Cloud Meeting pada google play dalam jangka waktu 3 bulan, yaitu dari bulan Maret – Mei 2020.
2. Data hasil review analisis pengguna Google meet dan Zoom Cloud Meeting, sudah disertakan dengan skor perolehan “BINTANG” sesuai hasil penilaian pengguna aplikasi dan dipisahkan menjadi 2 kelas (pertama, dengan skor bintang 1 sampai dengan 5 dan kedua, dengan skor bintang 1 dan bintang 5).
3. Tahapan pre-processing review komentar diproses dengan menggunakan tools Gata Framework Textmining yang dapat diakses melalui <http://www.gataframework.com/textmining/>, dengan menggunakan transform cases, tokenization, filtering, Stopword dan Stemming. Tahapan pre-processing juga dilakukan dengan menggunakan tools RapidMiner sehingga proses pembersihan data dilakukan secara bertingkat.
4. Pengujian pertama menggunakan model Algoritma *Naïve Bayes* dan *K-Nearest Neighbor* tanpa feature SMOTE dan PSO untuk melihat hasil awal.
5. Pengujian kedua menggunakan Synthetic Minority Over-sampling Technique (SMOTE) Upsampling, diterapkan untuk menyelesaikan masalah ketidakseimbangan kelas pada dataset.
6. Perhitungan dan pengujian akan dilakukan menggunakan 10 Fold Cross Validation serta menggunakan Algoritma *Naïve Bayes* dan *K-Nearest Neighbor* untuk mendapatkan klasifikasi dan akurasi yang terbaik dari kedua algoritma tersebut.
7. Hasil dari kedua algoritma tersebut, akan langsung diproses menggunakan optimasi *feature Particle Swarm Optimization (PSO)*, hal ini bertujuan untuk meningkatkan hasil nilai *Accuracy*, *Kappa* dan *AUC*.
8. Hasil dari penggunaan feature selection SMOTE Upsampling dan feature optimasi PSO diatas, kemudian dibandingkan hasilnya dengan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* tanpa menggunakan *feature selection SMOTE Upsampling* dan *Feature Otpimasi PSO* untuk mendapatkan hasil algoritma yang terbaik.
9. Setelah hasil nilai *Accuracy*, *Kappa* dan *AUC* didapatkan, maka akan diperoleh analisis dan kesimpulan hasil evaluasi dari penggunaan kedua algoritma tersebut.



Gambar 2.3. Kerangka Pemikiran

2.2 Tinjauan studi

2.2.1. Penelitian Terkait

Penelitian mengenai klasifikasi sentimen telah dilakukan oleh Agus darmawan dan Syamsiah [5], penelitian ini berjudul “Analisis Model Support

Vector Machine Textmining Pada Komentar Positif dan Negatif Untuk Review Perbandingan Whatsapp Vs BBM” dimana isinya membahas tentang review pengguna instant messaging. Sebagai pengguna smartphone, akhir-akhir ini kita dibuat pusing oleh munculnya beberapa layanan instant messaging lintas platform mobile yang kelihatannya dalam waktu singkat dapat menjadi sangat populer. Mulai dari Whatsapp dan BBM rasanya saat ini sedang bersaing ketat dalam memperebutkan perhatian para pengguna smartphone. Sebelum kita memutuskan untuk menggunakan instant messaging, ada hal dapat dipelajari dari hasil komentar review perbandingan Whatsapp dan BBM. Membaca komentar review tersebut secara keseluruhan dapat memakan waktu, namun jika hanya sedikit komentar review yang dibaca evaluasi akan menjadi bias. Dari beberapa teknik tersebut yang paling sering digunakan untuk klasifikasi data adalah *Support Vector Machines* (SVM). SVM sendiri memiliki kelebihan yaitu mampu mengidentifikasi dengan *hyperplane* terpisah yang memaksimalkan margin antara dua kelas yang berbeda dengan menggunakan data komentar *review* produk smartphone android dan blackberry. Penelitian ini menghasilkan klasifikasi teks dalam bentuk positif atau negatif dari review layanan instant messaging lintas platform mobile. Pengukuran berdasarkan akurasi *Support Vector Machines*, sedangkan pengukuran akurasi diukur dengan confusion matrix dan kurva ROC. Hasil penelitian menunjukkan akurasi *Support Vector Machines* adalah 71.00 %.

Abu Salam, Junta Zeniarja, Rima Septiyan Uswatun Khasanah [2], penelitian ini membahas tentang bagaimana mengelompokkan komentar-komentar yang ada di facebook page J&T ke dalam kelas positif dan kelas negatif. Sistem klasifikasi komentar diharapkan dapat membantu untuk mengetahui respon positif dan negatif dari pengguna facebook yang memberikan komentarnya. Sentiment analysis digunakan untuk mengetahui sikap seseorang dalam konteks dokumen. Sentiment analysis memiliki tahapan preprocessing yang terdiri dari *case folding*, *tokenizing*, *stopword removal*, *stemming*. Pembobotan kata yang digunakan adalah term frequency–invers document frequency dan perhitungan similaritasnya menggunakan cosine similarity kemudian menggunakan *K-Nearest Neighbor* sebagai metode klasifikasinya. Hasil yang didapatkan dari implementasi metode KNN ini cukup baik dengan uji coba sebanyak 6 kali. Rata-rata accuracy tertinggi adalah 79.21% sedangkan accuracy terendah adalah 70.3%.

Umi Rofiqoh, Rizal Setya Perdana, M. Ali Fauzi [3], penelitian ini membahas tentang Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode *Support Vector Machine dan Lexicon Based Features*, dimana Analisis sentimen adalah salah satu cabang penelitian dari *Text Mining* yang berguna untuk mengklasifikasi dokumen teks berupa opini berdasarkan sentimen. Dokumen teks yang digunakan dalam penelitian berasal Twitter tentang opini masyarakat mengenai penyedia layanan telekomunikasi seluler. Metode yang digunakan adalah *Support Vector Machine* dengan menggunakan *Lexicon Based Features* sebagai pembaharuan fiturnya selain memakai fitur TF-IDF. Data yang digunakan pada penelitian ini sebanyak 300 data yang dibagi menjadi dua jenis data dengan perbandingan 70% untuk data latih dan 30% untuk data uji. Hasil akurasi sistem yang diperoleh dari analisis sentimen dengan metode *Support Vector Machine dan Lexicon Based Features* sebesar 79% menggunakan nilai degree sebesar 2, nilai konstanta learning rate 0,0001, serta jumlah iterasi maksimum sebanyak 50 kali. Sedangkan sistem analisis sentimen tanpa menggunakan *Lexicon Based Features* menghasilkan tingkat akurasi sebesar 84% dengan nilai parameter yang sama. Sehingga hasilnya lebih baik tanpa menggunakan *Lexicon Based Features*.

Nurirwan Saputra, Teguh Bharata Adji, Adhistya Erna Permanasari (2015), penelitian yang membahas tentang Jokowi merupakan seorang tokoh masyarakat dengan jenjang karir yang sangat cepat, dan tidak luput dari pandangan masyarakat baik itu positif, netral maupun negatif. Data mengenai Jokowi yang berisikan komentar positif, netral dan negatif yang berasal dari media sosial dan blog politik diperlukan dalam menentukan langkah-langkah yang harus diambil oleh Jokowi untuk mendapatkan kepercayaan dari masyarakat. Selain itu data yang sudah didapat perlu dievaluasi untuk menunjukkan urgensi diimplementasikannya *Pre-Processing* terhadap data, yaitu normalisasi dan stemming. Penelitian ini menggunakan search techniques dalam pengambilan data, sehingga pengambilan data dilakukan dengan efektif dan efisien. Search techniques dalam penelitian ini menggunakan Boolean searching dengan operator “AND”. Data yang sudah didapat dilabeli positif, netral dan negatif oleh penulis kemudian dikoreksi oleh ahli bahasa. Setelah itu dilakukan preProcessing baik itu mengubah kata tidak baku menjadi baku atau biasa disebut normalisasi menggunakan kamus dan mencari akar kata yaitu stemming dengan bantuan aplikasi Sastrawi Master. Selanjutnya dilakukan

juga tokenisasi N-Gram, Unigram, Bigram, dan Trigram terhadap kalimat, kemudian menghilangkan kata-kata yang umum digunakan dan tidak mempunyai Informasi yang berharga pada suatu konteks atau biasa disebut *stopword removal*, dan mempertahankan *emoticon*, karena *emoticon* merupakan simbol yang menunjukkan ekspresi seseorang ke dalam tulisan. Akurasi yang terbaik dalam penelitian ini adalah dengan dilakukan normalisasi dan stemming pada data sebesar 89,2655% menggunakan metode SVM, dan kemudian data yang dinormalisasi saja sebesar 88,7006% menggunakan metode SVM. Dalam penelitian ini, tidak ada ujicoba terhadap data yang dilakukan stemming saja, dikarenakan tahap yang harus dilakukan dalam stemming adalah melakukan normalisasi terlebih dahulu terhadap data.

Muljono, Dian Putri Artanti , Abdul Syukur, Adi Prihandono, De Rosal I. Moses Setiadi [4], penelitian ini membahas tentang Situs online marketplace merupakan tempat belanja yang saat ini digemari masyarakat karena menawarkan berbagai kemudahan. Sebagian masyarakat puas dengan pelayanan yang diberikan situs online marketplace yang mereka pilih dan sebagian lagi tidak. Respon yang diberikan konsumen tentang pelayanan situs online marketplace biasanya diungkapkan melalui media sosial, salah satunya adalah Twitter. Pada penelitian ini dilakukan analisis sentimen terhadap postingan opini pelanggan online *marketplace* di Indonesia pada *Twitter* yang nantinya bisa digunakan untuk menentukan rating online *marketplace* agar masyarakat tidak salah pilih situs marketplace untuk berbelanja di masa depan. Langkah pertama melakukan koleksi data opini masyarakat di Twitter dari situs belanja online. Kemudian *dilakukan pre-processing* pada data yang meliputi *cleansing data, case folding, tokenizing, case normalization, stop word, convert negation dan stemming*. Selanjutnya dilakukan proses pemberian label (kelas) pada data tersebut yang dilakukan oleh ahli bahasa. Hasil dari proses clustering tercipta dua kelas data yaitu kelas data positif dan kelas data negatif dengan jumlah total 1200 data. Data yang sudah memiliki kelas data ini, digunakan sebagai data training untuk mesin pengklasifikasi, dalam riset ini menggunakan algoritma mesin pengklasifikasi *Naïve Bayes*. Terakhir, mengukur kinerja dari mesin pengklasifikasi menggunakan 10-fold cross validation. Hasil evaluasi menunjukkan rata-rata akurasi sebesar 93.33%.

Tabel 2.4. Penelitian Terkait

Penulis	Judul	Metode	Hasil
Agus darmawan, Syamsiah (2016)	Analisis Model <i>Support Vector Machine</i> Textmining Pada Komentar Positif Dan Negatif Untuk Review Perbandingan Whatsapp Vs BBM	<i>Support Vector Machine</i> dengan <i>confusion matrix</i>	Nilai Akurasi 71%
Abu Salam, Junta Zeniarja, Rima Septiyan Uswatun Khasanah (2018)	Analisis Sentimen Data Komentar Sosial Media Facebook Dengan <i>k-Nearest Neighbor</i> (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekspres Indonesia)	<i>k-Nearest Neighbor</i>	Hasil Rata-rata akurasi tertinggi 79.21% sedangkan akurasi terendah adalah 70.3%.
Umi Rofiqoh, Rizal Setya Perdana, M. Ali Fauzi (2017)	Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada <i>Twitter</i> Dengan Metode	<i>Support Vector Machine</i> dan <i>Lexicon Based Features</i>	SVM dengan <i>Lexion Based Feature</i> akurasi 79% sedangkan SVM tanpa <i>Lexion Based Feature</i> Nilai akurasi 84%

	<i>Support Vector Machine dan Lexicon Based Features</i>		
Nurirwan Saputra, Teguh Bharata Adji, Adhistya Erna Permanasari (2015)	Analisis Sentimen Data Presiden Jokowi Dengan <i>Preprocessing</i> Normalisasi Dan Stemming Menggunakan Metode <i>Naive Bayes</i> dan SVM	<i>Naive Bayes</i> dan <i>Support Vector Machine</i>	Nilai Akurasi SVM sebesar 89.26%
Muljono, Dian Putri Artanti , Abdul Syukur, Adi Prihandono, De Rosal I. Moses Setiadi (2018)	Analisis Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma <i>Naive Bayes</i>	<i>Naive Bayes</i>	Nilai akurasi 93.33%

2.3 Tinjauan Objek Penelitian

2.3.1. Zoom Cloud Meeting

Zoom adalah pemimpin dalam komunikasi video perusahaan modern, dengan platform cloud yang mudah dan andal untuk konferensi video dan audio, kolaborasi, obrolan, dan webinar di seluruh perangkat seluler, desktop, telepon, dan sistem ruang. Zoom Rooms adalah solusi ruang konferensi berbasis perangkat lunak asli yang digunakan di seluruh dunia dalam ruang dewan, konferensi, ngerumpi, dan pelatihan dan ruang kelas. Didirikan pada 2011, Zoom membantu bisnis dan organisasi menyatukan tim dalam lingkungan tanpa pertemuan untuk menyelesaikan lebih banyak hal. Zoom adalah perusahaan publik di Nasdaq berkantor pusat di San Jose, California.

Setiap orang dapat membuat dan berpartisipasi dalam *video conference* hingga 100 orang. Terhubung dengan siapa pun di ponsel dan tablet berbasis Android, perangkat seluler lain, Windows, Mac, Zoom Rooms, sistem ruang H.323 / SIP, dan telepon. Fitur utama dari Zoom Cloud Meeting adalah bisa menjangkau orang secara instan untuk mengirim pesan, file, gambar, tautan, dan gif dengan mudah, cepat menanggapi atau bereaksi terhadap percakapan berulir dengan emoji, membuat atau bergabung dengan saluran obrolan publik dan pribadi.

2.3.2. Google Meet

Google Meet (Hangouts Meet/Meet) adalah salah satu aplikasi atau software yang dapat dimanfaatkan untuk tetap produktif dalam bekerja meski dilakukan dari rumah. Dilansir dari Software Advice, perangkat lunak ini merupakan sebuah aplikasi konferensi percakapan video secara online. Ini adalah versi lain dari Google Hangouts yang dikhususkan untuk bisnis dari semua ukuran. Google Meet terintegrasi dengan G Suite, yang memungkinkan pengguna untuk dapat bergabung langsung dari Kalender atau undangan yang dikirim via email.

setiap orang dapat membuat dan berpartisipasi dalam rapat video berkualitas tinggi hingga 250 orang, dengan aman. Fitur utama:

1. Selenggarakan rapat video definisi tinggi tanpa batas.
2. Bertemu dengan aman - rapat video dienkripsi saat transit dan tindakan anti-penyalahgunaan proaktif dapat membuat rapat Anda tetap aman.
3. Akses mudah - cukup bagikan link dan tamu undangan dapat bergabung ke rapat dengan sekali klik dari browser web desktop atau aplikasi seluler Meet.
4. Bagikan layar Anda untuk mempresentasikan dokumen, slide, dan lainnya.
5. Hadiri rapat dengan teks yang diberdayakan oleh teknologi ucapan ke teks Google.

BAB III

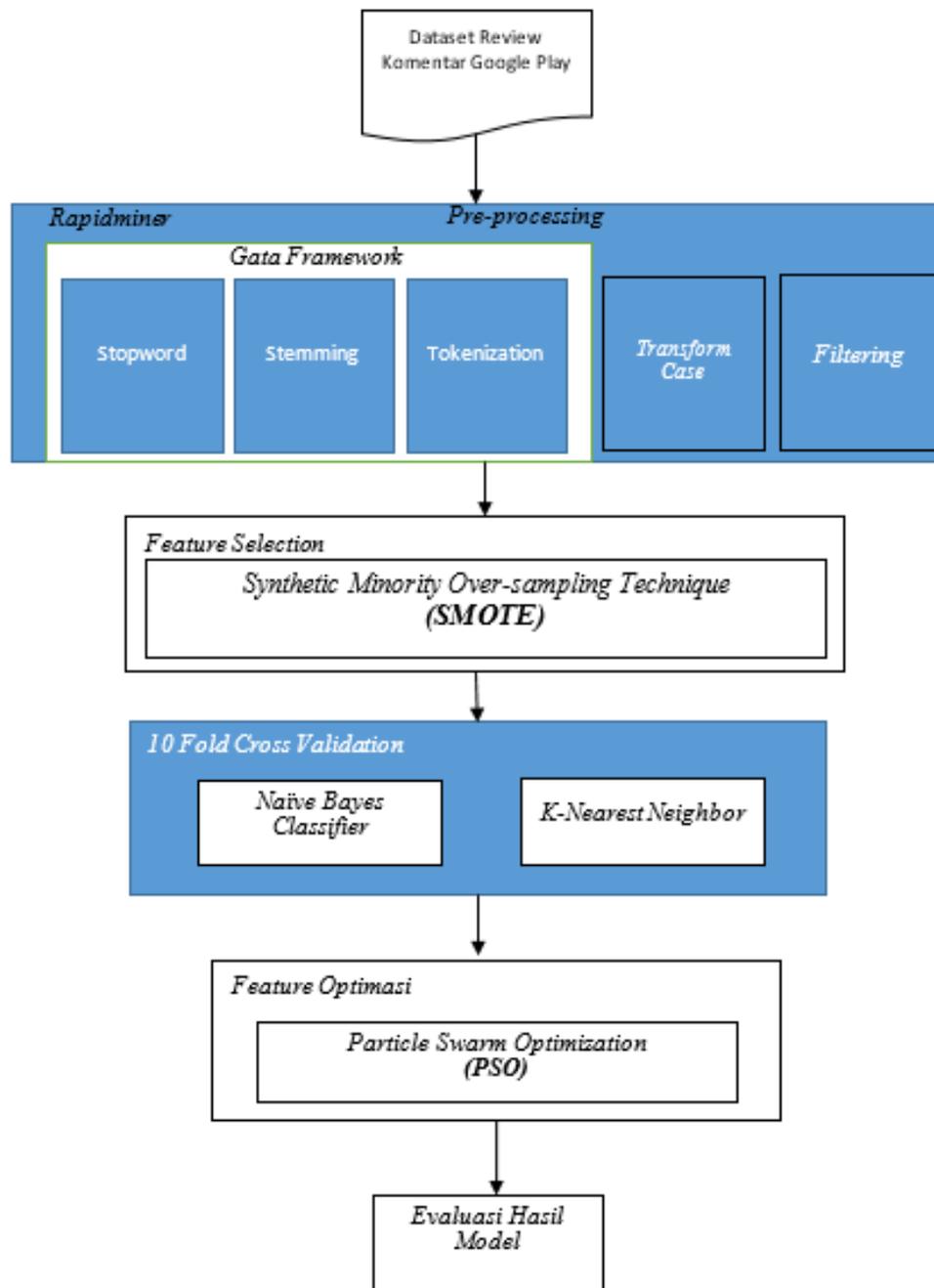
METODOLOGI PENELITIAN

Metodologi sendiri merujuk kepada alur pemikiran umum atau menyeluruh (general logic) dan gagasan teoritis (theoretic perspectives) suatu penelitian. Sedangkan kata 'metode' menunjuk pada teknik yang digunakan dalam penelitian seperti survey, wawancara dan observasi. Suatu penelitian yang baik senantiasa memperhatikan kesesuaian antara teknik yang digunakan dengan alur pemikiran umum serta gagasan teoritis [38].

Penelitian ini menggunakan data kuantitatif untuk menganalisis data dari teks berbahasa Indonesia berupa ulasan pengguna aplikasi Google Meet dan Zoom Cloud meeting yang didapat dari google play, data ulasan pengguna aplikasi Virtual meet baik Google Meet dan Zoom Cloud meeting yang sudah mendapatkan skor penilaian berupa banyaknya bintang yang diberikan oleh pengguna aplikasi tersebut. Dalam tahapan mencari pemecahan masalah ini akan dilakukan dengan kerangka pemikiran menggunakan model penelitian Knowledge Discovery in Database (KDD), dimana tahapannya meliputi:

1. Selection
2. Pre-Processing
3. Transformation
4. Data Mining
5. Interpretation/Evaluation

Masing-masing tahapan diatas, akan dijelaskan melalui gambar 3.1 dibawah ini untuk menjelaskan alur penelitian Analisis Review Pengguna Google Meet dan Zoom Cloud meeting Menggunakan Algoritma *Naïve Bayes* dan *K-Nearest Neighbor*.

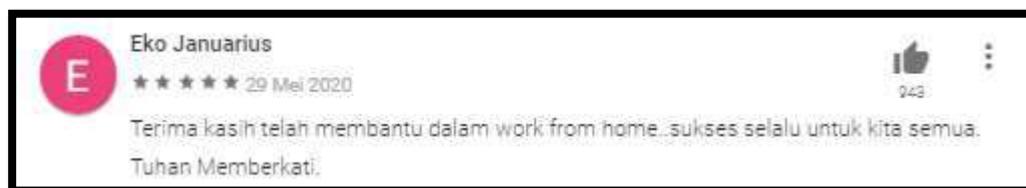


Gambar 3.1. Tahapan Penelitian

3.1 Metode Pengumpulan Data (Selection)

Untuk menunjang penelitian ini, dijelaskan tentang bagaimana dan dari mana data dalam penelitian ini didapatkan, pada tahapan ini juga dijelaskan mempersiapkan data yang diperlukan untuk dapat di proses agar data yang digunakan benar – benar sesuai dengan permasalahan yang akan dipecahkan. Data yang akan digunakan pada penelitian ini adalah review pengguna aplikasi Google

Meet dan Zoom Cloud Meeting pada google play yang memberikan komentar selama kurun waktu 3 bulan (Maret 2020 – Mei 2020), untuk menentukan status review komentar mengguna Google Meet dan Zoom Cloud Meeting yang ada di google play berdasarkan perolehan “Bintang”.



Gambar 3.2. Contoh review komentar pengguna Zoom Cloud meeting
Jumlah pencarian data review pengguna Google Meet dan Zoom Cloud Meeting pada *google play* dapat disampaikan pada tabel dibawah ini.

Tabel 3.1. Review Pengguna Google Meet dan Zoom Cloud Meeting

No	Review Pengguna	Jumlah Bintang					Total
		1	2	3	4	5	
1	Google Meet	152	24	35	39	254	504
2	Zoom Cloud Meeting	59	11	35	49	249	405
Jumlah							909

Sumber: Telah diolah kembali berdasarkan pencarian di *google play*

3.2 Teknik Analisis dan Pengujian Data (*Pre-processing*)

Pada tahap ini dilakukan klasifikasi teks atau sentimen dengan tahapan *preprocessing* agar teks yang memiliki isi yang tidak sempurna seperti data yang hilang, data yang tidak *valid* atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan. Data tersebut lebih baik dibuang karena keberadaannya bisa mengurangi mutu atau akurasi. Teks yang belum diolah biasanya memiliki karakteristik dimensi yang tinggi, terdapat noise pada data dan terdapat struktur teks yang tidak baik. Untuk itu, dalam pengolahan data awal, text mining harus melalui beberapa tahapan yang disebut dengan pre-processing. Tahapan pre-processing yang dapat dilakukan yaitu:

1. Tokenize

Tokenize merupakan proses untuk memisah-misahkan kata. Proses memotong setiap kata dalam teks dan mengubah huruf dalam dokumen menjadi huruf kecil. Hanya huruf yang diterima, sedangkan karakter khusus atau tanda baca akan dihilangkan.

2. Filter Tokens (By Length)

Menurut Langgeni bahwa filter token (by length) merupakan proses mengambil kata-kata penting dari hasil token. Dalam proses ini, kata-kata yang memiliki panjang tertentu akan dihapus.

3. Stopwords Removal

Menurut Langgeni bahwa filter stopwords removal adalah proses menghilangkan kata-kata yang sering muncul namun tidak memiliki pengaruh apapun dalam ekstraksi sentimen suatu review. Kata yang termasuk seperti kata penunjuk waktu, kata tanya.

4. Transform Cases

Dalam penulisan review komentar, pasti terdapat perbedaan bentuk huruf, tahapan ini merupakan proses merubah bentuk huruf menjadi huruf kecil (lower case) atau dapat disebut juga penyeragaman bentuk huruf.

5. Cleansing Data

Dilakukan untuk mengurangi noise pada data review komentar. Kata-kata yang tidak penting dihilangkan seperti URL, hashtag (#), username (@username), email, emoticon (:@, :*, :D), tanda baca seperti koma (,), titik (.) dan juga tanda baca lainnya. URL pada data juga dihapus pada tahap ini.

Kemudian tahapan pre-processing dilakukan sebanyak 2 kali, pertama dataset akan dilakukan tahap pre-processing menggunakan tools text-mining yang dibuat menggunakan PHP dan tersedia secara online melalui situs <http://www.gataframework.com/textmining/> dan tahapan kedua pre-processing dengan menggunakan tools RapidMiner.

3.3 Instrumen Penelitian (Transformation)

Instrumen dalam penelitian menggunakan model aplikasi Microsoft excel untuk mengambil data dan alat bantu untuk menghitung tingkat akurasi menggunakan RapidMiner versi 9.7. Dataset yang digunakan untuk pengujian model diperoleh dari review pengguna Google Meet dan Zoom cloud meeting pada google play. Hasil pengujian menggunakan perbandingan algoritma *Naïve Bayes* dan *K-NEAREST NEIGHBOR (KNN)*, namun sebelum dilakukan tahap pre-processing data training terlebih dahulu dilakukan proses Synthetic Minority Over-Sampling Technique (SMOTE) yaitu untuk mengatasi ketidakseimbangan jumlah data sehingga jumlah data review akan memiliki nilai yang sama untuk perolehan

bintangnya. Setelah penggunaan feature SMOTE, maka dilakukan pengujian dengan menggunakan algoritma *Naïve Bayes* dan *K-NEAREST NEIGHBOR (KNN)* dan untuk meningkatkan nilai akurasi maka akan ditambahkan feature Particle Swarm Optimization (PSO).

3.4 Model Yang Diusulkan (Data Mining)

Pada tahap ini dilakukan penentuan metode terhadap teknik data mining, alat bantu data mining, dan algoritma data mining yang akan diterapkan. Lalu selanjutnya adalah melakukan penerapan teknik dan algoritma data mining dengan alat bantu. Jika diperlukan penyesuaian data terhadap teknik data mining tertentu, dapat kembali ke tahap persiapan data. Tahap ini juga dapat disebut tahap learning karena pada tahap ini data training diklasifikasikan oleh model dan kemudian menghasilkan sejumlah aturan. Metode yang diusulkan adalah menerapkan feature selection yaitu Synthetic Minority Over-Sampling Technique (SMOTE) dan Particle Swarm Optimize (PSO) pada algoritma *Naïve Bayes* dan *K-NEAREST NEIGHBOR (KNN)*.

Peneliti akan melakukan perbandingan hasil akurasi dengan melakukan eksperimen terhadap model *Naïve Bayes*, *K-NEAREST NEIGHBOR (KNN)*, *Naïve Bayes+SMOTE Upsampling+PSO* dan *K-NEAREST NEIGHBOR (KNN)+SMOTE Upsampling+PSO*, untuk mengetahui model dan akurasi yang paling baik dalam menganalisis review pengguna Google Meet dan Zoom cloud meeting. Selain membandingkan algoritma model *Naïve Bayes* dan *K-NEAREST NEIGHBOR (KNN)*, penelitian ini juga akan membandingkan data review pengguna berdasarkan rating bintang dengan 2 metode yaitu metode pertama, data review yang mendapatkan rating bintang 1 sampai dengan bintang 5 dan metode kedua, data review yang mendapatkan rating bintang 1 dan bintang 5 saja.

3.5 Evaluasi dan Validasi Hasil

Pada tahapan ini diusulkan model untuk data review pengguna aplikasi Google Meet dan Zoom cloud meeting dengan menerapkan algoritma *Naïve Bayes* dan *K-NEAREST NEIGHBOR (KNN)* dengan menambahkan feature selection yaitu Synthetic Minority Over-Sampling Technique (SMOTE) dan Particle Swarm Optimize (PSO). Dalam penelitian menggunakan validasi standar yaitu 10 fold cross-validation dimana proses ini membagi data secara acak ke dalam 10 bagian.

Proses pengujian dimulai dengan pembentukan model dengan data pada bagian pertama. Model yang terbentuk akan diujikan pada 9 bagian data sisanya.

Proses yang dilakukan setelah melakukan pengujian yaitu mengukur performance dari algoritma klasifikasi text mining yang dipakai. Dalam penelitian ini performance diukur menggunakan nilai Accuracy, Kappa dan AUC.

BAB IV
HASIL PENELITIAN DAN PEMBAHASAN

4.1 Pengumpulan Data

Pada penelitian ini, review pengguna Google Meet dan Zoom Cloud Meeting data diambil dari google play, dimana review pengguna yang diambil sejumlah 909 data (perolehan skor penilaian bintang 1 sampai dengan bintang 5), untuk review pengguna Google Meet berjumlah 504 dan review pengguna Zoom Cloud Meeting berjumlah 405, kemudian dari 909 data tersebut akan dilakukan pengujian kedua, untuk membandingkan perolehan skor penilaian bintang 1 dan bintang 5 dimana jumlahnya sekitar 714 data. Dari review pengguna tersebut, kemudian dikelompokkan berdasarkan skor penilaian bintang yang diberikan oleh pengguna Google Meet dan Zoom Cloud Meeting, dimana untuk gambarannya seperti tabel berikut ini:

Tabel 4.1. Jumlah Data Review Pengguna Bintang 1 – 5

No	Review Pengguna	Jumlah Bintang					Total
		1	2	3	4	5	
1	Google Meet	152	24	35	39	254	504
2	Zoom Cloud Meeting	59	11	35	49	249	405
Jumlah							909

Sumber: Telah diolah kembali berdasarkan pencarian di *google play*

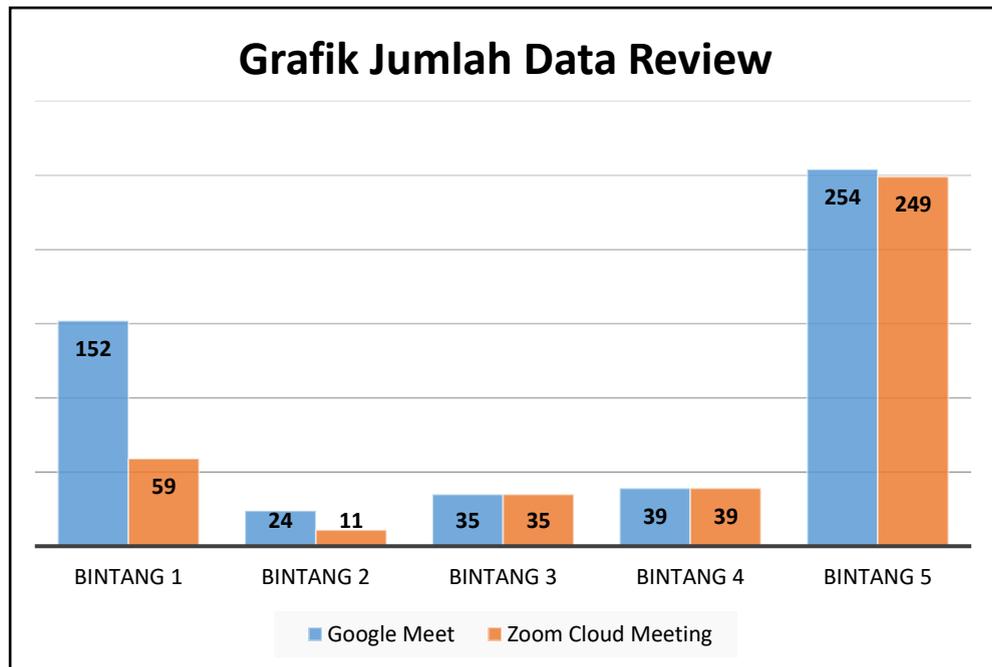
Tabel 4.2. Jumlah Data Review Pengguna Bintang 1 dan 5

No	Review Pengguna	Jumlah Bintang		
		1	5	Total
1	Google Meet	152	254	406
2	Zoom Cloud Meeting	59	249	308
Jumlah				714

Sumber: Telah diolah kembali berdasarkan pencarian di *google play*

Dari data diatas, jika dibuatkan grafik berdasarkan perolehan bintang dimana untuk gambaran seperti tabel 4.3. berikut ini:

Tabel 4.3. Grafik Jumlah Data Review Aplikasi Google Meet dan Zoom Cloud Meeting



Sumber: Telah diolah kembali berdasarkan pencarian di *google play*

Dalam mengolah data untuk mendapatkan model yang sesuai dengan kasus penelitian ini yaitu Review Analisis Pengguna Google Meet dan Zoom Cloud Meeting menggunakan Algoritma *Naïve Bayes* dan *K-Nearest Neighbor* digunakanlah sebuah tools Rapidminer versi 9.7. Karena penelitian ini termasuk bagian dari text-mining maka ada tahap yang harus dilakukan terlebih dulu sebelum bisa didapatkan sebuah model yang baik pada studi kasus analisis komentar pengguna Google Meet dan Zoom Cloud Meeting pada google play.

4.2 Pengolahan data

Pembahasan pada tahapan ini adalah proses awal pengolahan dataset sebelum dapat diproses untuk klasifikasi dengan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* yang dioptimalkan dengan menggunakan feature Synthetic Minority Over-Sampling Technique (SMOTE) dan *Particle Swarm Optimization* (PSO). Penelitian ini menggunakan beberapa tahapan *pre-processing* untuk dataset teks komentar, berikut ini adalah tahapan tersebut:

1. Case Folding
2. Tokenizing

3. Filtering
4. Stemming

Tahapan-tahapan pre-processing teks komentar diatas diproses dengan menggunakan tools Gata Framework Textmining yang dapat diakses melalui www.gataframework.com/textmining, untuk menggunakan tools tersebut dapat dalam single teks yang dapat langsung di masukan pada kolom yang tersedia di halaman tools tersebut, namun juga dapat dalam bentuk file excel yang kemudian diunggah. Berikut ini adalah tampilan halaman pre-processing teks dengan menggunakan tools Gata Framework Textmining, pada gambar 4.1.



Gambar 4.1. Tampilan tools Gata Framework Textmining

Gata Framework merupakan alternatif dalam pre-processing teks berbahasa Indonesia yang dikombinasikan dengan aplikasi RapidMiner untuk memproses kata-kata dalam bahasa Indonesia, hal ini dikarenakan dalam aplikasi RapidMiner sudah ada fasilitas kamus untuk mengubah akronim, dan stopwords, tetapi masih terbatas pada bahasa Inggris, Cina, dan Arab, sedangkan untuk bahasa Indonesia masih belum tersedia. Dari hasil pre-processing dengan menggunakan Gata Framework, maka dataset akan dilakukan pre-processing lagi dengan menggunakan tools RapidMiner untuk membersihkan data agar lebih baik lagi hasilnya.

Berikut hasil penggunaan tools Gata Framework berupa file Ms. Excel yang akan diolah kedalam tools RapidMiner seperti gambar 4.2 dan gambar 4.3.

A	B	C	D	E	F
1	no	rating	Tokenization: Regexp	Indonesian Stemming	Indonesian Stop word removal
1	Maaf ini ada apa ya kenapa saya ngratal aplikasi tidak berunduh padahal kan saya udah pakai akunnya. App ada yang salah?	5	Maaf ini ada apa ya kenapa saya ngratal aplikasi tidak berunduh padahal kan saya udah pakai akunnya app ada yang salah?	Maaf ini ada apa ya kenapa saya ngratal aplikasi tidak unduh padahal kan saya udah pakai aku app ada yang salah?	Maaf ngratal aplikasi unduh pakai salah?
2	Aplikasi Google kelas ini sangat bermanfaat dan membantu dalam memberikan materi kuliah, bimbingan tugas akhir mahasiswa dan membuka ruang konsultasi bagi Dosen pembimbing di kampus. Apalagi kita di Indonesia sedang dalam situasi Darurat bencana kesehatan masyarakat dampak pandemi Covid-19 dibatasi akses internet di rumah.	5	Aplikasi google kelas ini sangat bermanfaat dan membantu dalam memberikan materi kuliah bimbingan tugas akhir mahasiswa dan membuka ruang konsultasi bagi dosen pembimbing di kampus. Apalagi kita di Indonesia sedang dalam situasi darurat bencana kesehatan masyarakat dampak pandemi covid dibatasi akses internet di rumah.	Aplikasi google kelas ini sangat manfaat dan bantu dalam berik materi kuliah bimbingan tugas akhir mahasiswa dan buka ruang konsultasi bagi dosen bimbing di kampus. Apalagi kita di Indonesia sedang dalam situasi darurat bencana sehat masyarakat dampak pandemi covid iaku physical distanse di wilayah.	Aplikasi google kelas manfaat bantu berik materi kuliah bimbingan tugas mahasiswa. Buka ruang konsultasi dosen bimbingan kampus Indonesia situasi darurat. Bencana sehat masyarakat dampak pandemi covid iaku physical distanse di wilayah.
3	Saya sangat tertarik dengan aplikasi Google classroom ini apalagi di saat sekarang ini ilmu saya masih minim sekali tentang ini, mudah-mudahan kita bisa membawa anak didik kita dengan pembelajaran online ini, walaupun kita tidak dapat bertatap muka langsung dengan siswa, tapi pembelajaran tetap bisa dilakukan dengan baik juga bisa kalau bisa dikasih fitur unduh untuk setiap file yang dikirim bapak ibu guru.	5	Saya sangat tertarik dengan aplikasi google classroom ini apalagi di saat sekarang ini ilmu saya masih minim sekali tentang ini mudah-mudahan kita bisa membawa anak didik kita dengan pembelajaran online ini walaupun kita tidak dapat bertatap muka langsung dengan siswa tapi pembelajaran tetap bisa dilakukan dengan baik juga bisa kalau bisa dikasih fitur unduh untuk setiap file yang dikirim bapak ibu guru.	Saya sangat tarik dengan aplikasi google classroom ini apalagi saat sekarang ini ilmu saya masih minim kali tentang ini mudah-mudah kita bisa bawa anak didik kita dengan pembelajaran online ini walaupun kita tidak dapat tatap muka langsung dengan siswa tapi pembelajaran tetap langsung dengan baik juga bisa manfaat teknologi.	tarik aplikasi google classroom ilmu minim kali mudah-mudah bawa anak didik pembelajaran online tatap muka langsung siswa pembelajaran langsung manfaat teknologi.
4	Jadi saya mencoba membuka kelas saya, tetapi dikatakan bukan kelas saya dan kemudian segera menghapusnya dari pandangan saya. Saya memeriksa ruang kelas di komputer saya untuk melihat apakah ada sesuatu yang salah dengan akun saya, tetapi semua kelas saya ada di sana, jadi saya cukup yakin bahwa aplikasi tidak berfungsi. Ini hanya ketidaknyamanan utama, karena saya tidak dapat menggunakan ponsel saya.	3	Jadi saya mencoba membuka kelas saya tetapi dikatakan bukan kelas saya dan kemudian segera menghapusnya dari pandangan saya. Saya memeriksa ruang kelas di komputer saya untuk melihat apakah ada sesuatu yang salah dengan akun saya tetapi semua kelas saya ada di sana jadi saya cukup yakin bahwa aplikasi tidak berfungsi ini hanya ketidaknyamanan utama karena saya tidak dapat menggunakan ponsel saya.	Jadi saya coba buka kelas saya tetapi katak bukan kelas saya dan kemudian segera hapus dari pandang saya saya periksa ruang kelas di komputer saya untuk melihat apakah ada suatu yang salah dengan akun saya tetapi semua kelas saya ada di sana jadi saya cukup yakin bahwa aplikasi tidak fungsi ini hanya ketidaknyamanan utama karena saya tidak dapat guna ponsel saya untuk mendaftarkan hal hal seperti kerja rumah.	coba buka kelas katak kelas hapus pandang saya periksa ruang kelas komputer lihat salah akun kelas saya aplikasi fungsi ketidaknyamanan utama ponsel mendaftarkan hal kerja rumah informasi bentuk milik akses komputer.

Gambar 4.2. Hasil pre-processing dengan menggunakan tools Gataframework untuk pengguna Google Meet

A	B	C	D	E	F
1	no	rating	Tokenization: Regexp	Indonesian Stemming	Indonesian Stop word removal
1	Saya suka sekali, karena sangat membantu meetings jarak jauh, apalagi di saat korona melanda di Indonesia, semua orang tidak boleh keluar rumah.	5	Saya suka sekali karena sangat membantu meetings jarak jauh apalagi di saat korona melanda di Indonesia semua orang tidak boleh keluar rumah.	Saya suka kali karena sangat bantu meetings jarak jauh apalagi saat korona lande di Indonesia semua orang tidak boleh keluar rumah.	Suka kali bantu meetings jarak korona lande Indonesia orang rumah.
2	Bisa belajar bareng	5	Bisa belajar bareng	bisa ajar bareng	ajar bareng
3	Saya baru akan mencoba untuk belajar pakai Zoom ya Terima kasih	5	Saya baru akan mencoba untuk belajar pakai zoom ya terima kasih	Saya baru akan coba untk ajar pakai zoom ya terima kasih	coba ajar pakai: zoom ya terima kasih
4	Tidak bisa masuk app nya	1	Tidak bisa masuk app nya	tidak bisa masuk app nya	masuk app
5	Terima kasih zoom	5	Terima kasih zoom	terima kasih zoom	terima kasih zoom
7	Di download nya lama banget ya dari tadi	2	Di download nya lama banget ya dari tadi	di download nya lama banget ya dari tadi	download banget
7	bagus mas lur tapi belum banyak teman kantor yang gabung disini	5	bagus mas lur tapi belum banyak teman kantor yang gabung disini	bagus mas lur tapi belum banyak teman kantor yang gabung sini	bagus mas lur teman kantor gabung
8	Mahal amat pakai aplikasi nya	1	mahal amat pakai aplikasi nya	mahal amat pakai aplikasi nya	mahal pakai aplikasi
9	lumayan gus, jadi lebih hemat biaya waktu dan tenaga lebih praktis cuma tidak bisa makan bareng bareng dalam suasana yg sama	4	lumayan gus jadi lebih hemat biaya waktu dan tenaga lebih praktis cuma tidak bisa makan bareng bareng dalam suasana yg sama	lumayan gus jadi lebih hemat biaya waktu dan tenaga lebih praktis cuma tidak bisa makan bareng bareng dalam suasana yg sama	lumayan gus hemat biaya tenaga praktis. makan bareng bareng suasana
10	Manfaat bisa ketemu keluarga	4	manfaat bisa ketemu keluarga	manfaat bisa ketemu keluarga	manfaat ketemu keluarga
11	Saya sudah instal terus cara buka untuk memulai di mana ya memulai nya	4	Saya sudah instal terus cara buka untuk memulai di mana ya memulai nya	Saya sudah instal terus cara buka untuk mulai di mana ya mulai	instal buka
12	Aplikasi nya mantap	5	Aplikasi nya mantap	aplikasi mantap	aplikasi mantap
13	Aplikasi ini bisa buat laptop kan? tapi kenapa saya ga bisa sudah terinstal tapi ga mau dibuka kenapa dan kasih tau caranya tolong	2	Aplikasi ini bisa buat laptop kan tapi kenapa saya ga bisa sudah terinstal tapi ga mau dibuka kenapa dan kasih tau caranya tolong	Aplikasi ini bisa buat laptop kan tapi kenapa saya ga bisa sudah terinstal tapi ga mau buka kenapa dan kasih tau cara tolong	aplikasi laptop terinstal buka kasih tau tolong
14	layak sekali untuk kegiatan pembelajaran	5	layak sekali untuk kegiatan pembelajaran	layak kali untuk kiat pembelajaran dan rakor	layak kali untuk pembelajaran rakor hemat

Gambar 4.3. Hasil pre-processing dengan menggunakan tools Gataframework untuk pengguna Zoom Cloud Meeting

Dari tabel diatas, akan dirapikan formatnya dimana yang akan digunakan adalah kolom terakhir “Indonesian stop word removal”, seperti gambar 4.4 dan gambar 4.5. file yang sudah selesai di pre-processing dengan Gata Framework, maka akan disimpan dalam bentuk Ms. Excel untuk nantinya akan dilakukan pengujian dan penghitungan dengan tools RapiMiner.

no	text	rating
1	maaf instal aplikasi unduh pakai salah	1
2	coba buka kelas kata kelas hapus pandang saya periksa ruang kelas komputer lihat salah akun kelas saya aplikasi fungsi tidak nyaman utama port	1
3	aplikasi error masuk	1
4	bug kelas tiba logout	1
5	guru beri masuk aneh woi	1
6	masuk logo aplikasi lambang ikan	1
7	tiba data kelas hapus payah	1
8	kasih bintang pb mayor	1
9	masuk akun google	1
10	kelas	1
11	bug kelas log in tolong logo ikan lambang ikan	1
12	kelas kelas nampak aplikasi buka web buka	1
13	google classroom nya bug buka tolong bug nya cepat tugas kumpul	1
14	pas lihat tugas sesuai kelas daftar kelas hapus masuk pakai akun	1
15	aplikasi habis perbarui bug kelas	1
16	pas perbarui masuk akun classroom refresh masuk mohon karna sangkut tugas	1
17	download masuk aplikasi	1
18	masuk akun google saya akun google masuk mohon bantu	1
19	perbaiki bug	1
20	abis perbarui bug masuk tolong	1
21	habis perbarui masuk diinstall ulang tetap error masuk	1
22	perbarui logout class masuk	1
23	simbol ikan logo masuk	1

Gambar 4.4. Dataset Google Meet yang akan digunakan dalam RapidMiner

no	text	rating
1	suka kali bantu napat jarak korona lante Indonesia orang rumah	5
2	nyar bareng	5
3	coba ajak pakai zoom terima kasih	5
4	masuk aplikasi	1
5	terima kasih zoom	5
6	download banget	3
7	logus mas lur teman kantor gabung	5
8	mahai pakai aplikasi	1
9	lumayan guai hemat biaya tenaga praktis makan barung barang susane	4
10	manfaat ketemu keluarga	4
11	natal buke	4
12	aplikasi mantap	5
13	aplikasi laptop terinstal buka kasih tau tolong	2
14	layak kali buat pembelajaran rakor hemat biaya efisien tepat cocok aplikasi an darurat	5
15	coba	3
16	esal/mau/leluon/alamat pagi	5
17	suka	5
18	manfaat	4
19	faham	4
20	logus	3
21	download lancar	5
22	ajar presentasi seminar network kareeeeen	3
23	tes	5

Gambar 4.5. Dataset Zoom Cloud Meeting yang akan digunakan dalam RapidMiner

4.3. Metode Yang Diusulkan

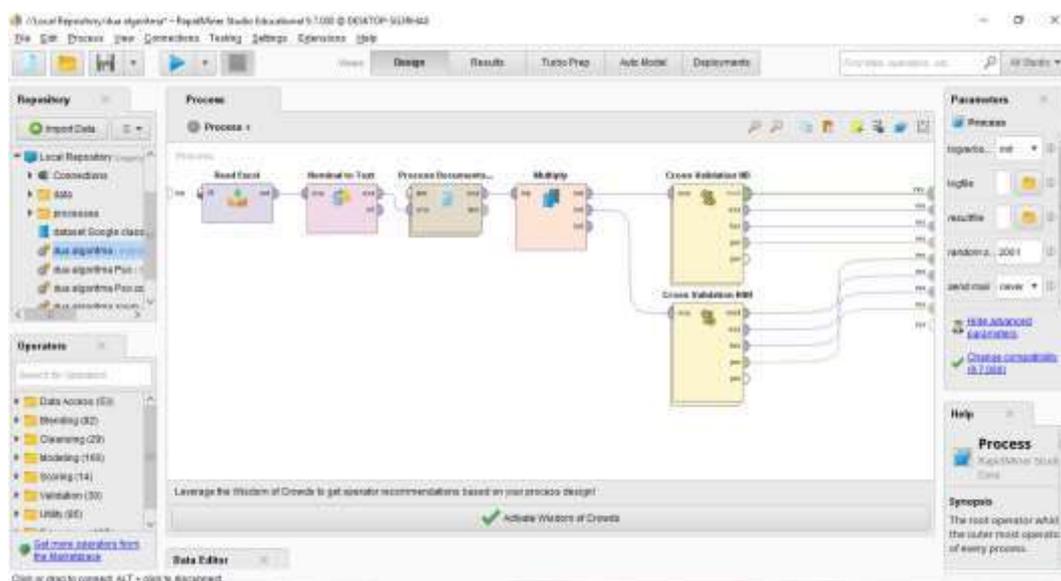
Pada tahap ini adalah membuat model dengan menggunakan algoritma klasifikasi untuk dataset yang sudah melalui tahap pre-processing. Tahap ini menggunakan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dengan menggunakan tools rapidminer versi 9.7 untuk mengolah dataset. Langkah dari proses ini adalah dataset akan diunggah ke dalam *tools* dengan menggunakan file excel yang akan di proses dengan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* untuk mendapatkan hasil awal dari masing-masing.

Jika hasil perhitungan awal dari penggunaan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* menunjukkan nilai *accuracy* dan *kappa* masih kurang bagus,

maka langkah selanjutnya akan ditambahkan penggunaan *feature selection* untuk meningkatkan nilai *accuracy* maupun kappa.

4.4. Evaluasi model klasifikasi

Tahapan evaluasi bertujuan untuk menentukan hasil kegunaan dari model yang telah dibuat pada langkah sebelumnya. Untuk evaluasi digunakan 10 fold-cross validation. Pada pengujian ini, data yang digunakan adalah data yang sudah melalui pre-processing. data tersebut diambil dari operator read excel, hal ini dilakukan karena dataset disimpan dalam format excel. Proses *documents from file* untuk mengkonversi file menjadi dokumen. Proses validasi terdiri dari data pelatihan dan data pengujian, untuk pengujian pertama adalah menjalankan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* dengan desain model seperti gambar 4.6. pada pengujian ini dilakukan dengan 2 metode, yaitu metode pertama menggunakan data analisa *review* pengguna dengan perolehan bintang 1 sampai 5 (5 kelas) dan metode kedua menggunakan data analisa review pengguna dengan perolehan bintang 1 dan bintang 5 (2 kelas), sehingga nantinya dapat dibandingkan perolehan nilai hasil akurasi.



Gambar 4.6. Desain Model menggunakan Algoritma NB dan K-NN

4.4.1 Pengujian pertama (5 kelas)

Untuk pengujian pertama, menggunakan data review pengguna yang memberikan skor penilaian bintang 1 - bintang 5, dimana setelah desain model ini dijalankan, maka akan tampak hasil nilai *accuracy* dan kappa yang dihasilkan dari penggunaan algoritma *Naïve Bayes* dan *k-nearest neighbor* baik Google Meet maupun Zoom Cloud Meeting.

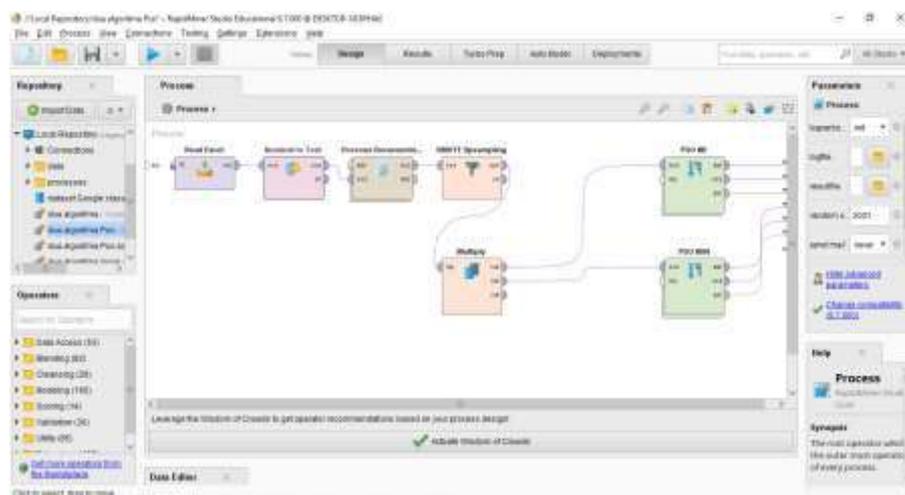
Tabel 4.4. Perbandingan Nilai *Accuracy* dan *Kappa* dengan *NB* dan *k-NN* (metode pertama: 5 Kelas, Bintang 1 – Bintang 5)

Algoritma	akurasi		Kappa	
	Google Meet	Zoom Cloud Meeting	Google Meet	Zoom Cloud Meeting
<i>Naïve Bayes</i>	34,52 %	33,34 %	0,123	0,056
KNN	67,25 %	64,95 %	0,425	0,196

Dari hasil tabel 4.4 diatas, maka hasil pengujian menunjukkan nilai *Accuracy* dan *Kappa* masih rendah, sehingga perlu ditambahkan metode penambahan feature optimasi untuk meningkatkan nilai *Accuracy* dan *Kappa*. Beberapa *feature* optimasi yang akan diuji cobakan antara lain:

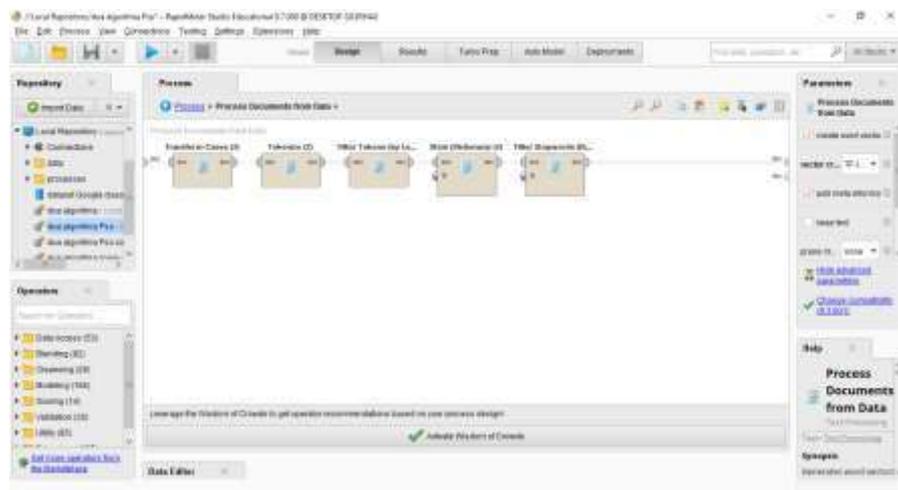
1. Penggunaan *feature Synthetic Minority Over-Sampling Technique* (SMOTE), karena merupakan salah satu algoritma khusus yang efektif untuk membangun kembali keseimbangan kelas dalam kasus *oversampling*. SMOTE menyeimbangkan kumpulan data dengan mengumpulkan data minoritas secara sintetis di ruang masukan berdasarkan informasi lingkungan mereka.
2. Penggunaan *feature Particle Swarm Optimization* (PSO), karena *Particle Swarm Optimization* (PSO) banyak digunakan untuk memecahkan masalah optimasi serta sebagai masalah seleksi fitur.

Karena pengujian pertama belum menunjukkan hasil yang signifikan, maka desain model akan berubah dengan menggunakan 2 *feature* optimasi untuk meningkatkan nilai *Accuracy* dan *Kappa* yaitu *SMOTE Upsampling* dan *PSO*, maka untuk desain model klasifikasi algoritma *Naïve Bayes* dan *K-Nearest Neighbor* akan mengalami perubahan modelnya, seperti gambar 4.7 dibawah ini.

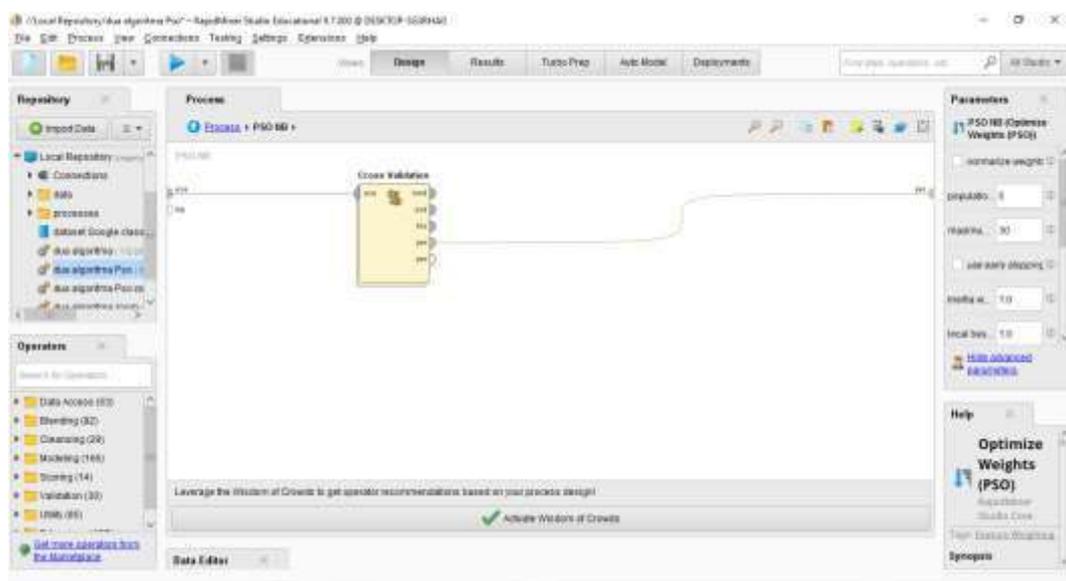


Gambar 4.7. Desain Model Algoritma *NB* dan *k-NN* dengan *feature SMOTE* dan *PSO*Gambar 4.8. Desain Evaluasi *10 Fold Cross-Validation* untuk algoritma *Naïve Bayes*Gambar 4.9. Desain Evaluasi *10 Fold Cross-Validation* untuk algoritma *k-Nearest Neighbor*

Untuk process documents from Data didalam tools Rapidminer 9.3, peneliti melakukan proses *pre-processing* untuk kedua kalinya walaupun sebelumnya sudah dilakukan dengan menggunakan tools gataframework text mining dengan alur tahapan *pre-processing* seperti gambar 4.10



Gambar 4.10. Desain Evaluasi *Process Documents*



Gambar 4.11. Desain model *PSO* Algoritma *Naïve Bayes*

Desain model Algoritma *Naïve Bayes* dan *K-Nearest Neighbour* diatas diterapkan untuk dataset review pengguna Zoom Cloud Meeting dan Google Meet, kemudian untuk langkah berikutnya adalah menjalankan proses eksekusi didalam *tools* Rapidminer 9.7 untuk mendapatkan hasil nilai *Accuracy* dan *Kappa* baik untuk review pengguna Google Meet maupun Zoom Cloud Meeting seperti tabel dibawah ini:

Tabel 4.5. Hasil Pengujian Algoritma *Naïve Bayes* dan k-Nearest Neighbor dengan feature SMOTE Upsampling dan PSO (metode pertama: 5 Kelas, Bintang 1 - Bintang 5)

Algoritma	akurasi	Kappa
-----------	---------	-------

	Google Meet	Zoom Cloud Meeting	Google Meet	Zoom Cloud Meeting
<i>Naïve Bayes</i> + smote+PSO	58,87 %	61,60 %	0,440	0,454
KNN + Smote+ PSO	73,84%	77,92%	0,611	0,644

4.4.2. Pengujian Kedua (2 kelas)

Untuk pengujian kedua ini, menggunakan data review pengguna yang memberikan nilai skor penilaian bintang 1 dan bintang 5 saja, sehingga hasil yang didapatkan seperti pada tabel 4.6 dibawah ini.

Tabel 4.6. Perbandingan Nilai *Accuracy* dan *AUC* dengan Algoritma *NB* dan *k-NN* (metode kedua: 2 Kelas, Bintang 1 dan Bintang 5)

Algoritma	akurasi		AUC	
	Google Meet	Zoom Cloud Meeting	Google Meet	Zoom Cloud Meeting
<i>Naïve Bayes</i>	66,74 %	72,42%	0,762	0,705
KNN	83,24 %	79,51%	0,851	0,696

Dari hasil pengujian algoritma *Naïve Bayes* dan *K-Nearest Neighbor* menghasilkan nilai *Accuracy* yang cukup baik yaitu diatas 60%. Jika pengujian kedua ini, menggunakan feature optimasi (SMOTE dan PSO) seperti yang dilakukan pada pengujian pertama, maka Algoritma *K-Nearest Neighbor* menghasilkan nilai *Accuracy* dan *AUC* mencapai 90% seperti pada tabel 4.7 dibawah ini.

Tabel 4.7. Perbandingan Nilai *Accuracy* dan *AUC* dengan Algoritma *NB* dan *k-NN* dengan feature SMOTE Upsampling dan PSO (metode kedua: 2 Kelas, Bintang 1 dan Bintang 5)

Algoritma	akurasi		AUC	
	Google Meet	Zoom Cloud Meeting	Google Meet	Zoom Cloud Meeting
<i>Naïve Bayes</i> + smote+PSO	79,33%	85,76 %	0,852	0,961
KNN + Smote+ PSO	83,46 %	91,56 %	0,911	0,955

4.5 Evaluasi dan Hasil Pengujian

Hasil dari pengujian model yang dilakukan adalah mengklasifikasikan *review* pengguna Google Meet dan Zoom Cloud meeting pada google play berdasarkan perolehan bintang sebagai dataset menggunakan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* berbasis *Synthetic Minority Over-sampling Technique* (SMOTE) dan Particle Swarm Optimization (PSO), berikut ini adalah hasil pengujian pada tools rapidminer sebagai berikut:

4.5.1. Hasil Pengujian Metode Pertama – 5 Kelas

4.5.1.1 Hasil Pengujian Google Meet dengan Model *Naïve Bayes* (NB)

Untuk hasil analisis review pengguna Google Meet dengan menggunakan algoritma *Naïve Bayes* sebagai berikut:

Accuracy : 34,52 %

Kappa : 0,123

Tabel 4.8. Hasil review pengguna Google Meet dengan Algoritma *Naïve Bayes*

	true 1	true 5	true 3	true 4	true 2	class precision
pred. 1	76	50	15	10	12	46.63%
pred. 5	16	91	5	13	5	70.00%
pred. 3	17	33	2	7	5	3.12%
pred. 4	9	44	3	5	2	7.94%
pred. 2	34	36	10	4	0	0.00%
class recall	50.00%	35.83%	5.71%	12.82%	0.00%	

$$\text{Acc (Accuracy)} = \frac{174}{504} = 0,345$$

Hasil akurasi yang didapatkan sejumlah 34.52%, dari tabel 4.8 maka hasil class precision dapat disimpulkan sebagai berikut:

- Perolehan Bintang 5 (SANGAT BAGUS) menduduki posisi pertama
- Perolehan Bintang 1 (SANGAT KURANG) menduduki posisi kedua
- Perolehan Bintang 4 (BAGUS) menduduki posisi ketiga
- Perolehan Bintang 3 (CUKUP) menduduki posisi keempat
- Perolehan Bintang 2 (KURANG) menduduki posisi kelima

Sehingga dari hasil analisis diatas, maka untuk review pengguna Google Meet dengan menggunakan Algoritma *Naïve Bayes* termasuk kategori “Sangat Bagus”.

4.5.1.2 Hasil Pengujian Google Meet dengan Model *k-Nearest Neighbor* (*k-NN*)

Untuk hasil analisis review pengguna Google Meet dengan menggunakan algoritma *k-Nearest Neighbor (k-NN)* sebagai berikut:

Accuracy : 67,25%

Kappa : 0,425

Tabel 4.9. Hasil review pengguna Google Meet dengan Algoritma *Naïve bayes*

	true 1	true 5	true 3	true 4	true 2	class precision
pred. 1	105	21	19	7	16	62.50%
pred. 5	43	233	16	31	6	70.82%
pred. 3	1	0	0	1	1	0.00%
pred. 4	0	0	0	0	0	0.00%
pred. 2	3	0	0	0	1	25.00%
class recall	69.08%	91.73%	0.00%	0.00%	4.17%	

$$\text{Acc (Accuracy)} = \frac{339}{504} = 0,672$$

Hasil akurasi yang didapatkan sejumlah 67.25 %, dari tabel 4.9 maka hasil class precision dapat disimpulkan sebagai berikut:

- Perolehan Bintang 5 (SANGAT BAGUS) menduduki posisi pertama
- Perolehan Bintang 1 (SANGAT KURANG) menduduki posisi kedua
- Perolehan Bintang 2 (Kurang) menduduki posisi ketiga
- Perolehan Bintang 3 (CUKUP) menduduki posisi keempat
- Perolehan Bintang 4 (BAGUS) menduduki posisi kelima

Sehingga dari hasil analisis diatas, maka untuk review pengguna Google Meet dengan menggunakan Algoritma *k-Nearest Neighbor (k-NN)* termasuk kategori “Sangat Bagus”.

4.5.1.3 Hasil Pengujian Zoom Cloud Meeting dengan Model *Naïve Bayes (NB)*

Untuk hasil analisis review pengguna Zoom Cloud Meeting dengan menggunakan algoritma *Naïve Bayes* sebagai berikut:

Accuracy : 33,34 %

Kappa : 0,056

Tabel 4.10. Hasil review pengguna Zoom Cloud Meeting dengan Algoritma *Naïve Bayes*

	true 5	true 1	true 2	true 4	true 3	class precision
pred. 5	96	14	3	21	18	63.16%
pred. 1	25	19	1	9	3	33.33%
pred. 2	30	7	2	4	3	4.35%

pred. 4	63	10	4	9	4	10.00%
pred. 3	35	9	1	6	9	15.00%
class recall	38.55%	32.20%	18.18%	18.37%	24.32%	

$$\text{Acc (Accuracy)} = \frac{135}{405} = 0,333$$

Hasil akurasi yang didapatkan sejumlah 33.34 %, dari tabel 4.10 maka hasil class precision dapat disimpulkan sebagai berikut:

- Perolehan Bintang 5 (SANGAT BAGUS) menduduki posisi pertama
- Perolehan Bintang 1 (SANGAT KURANG) menduduki posisi kedua
- Perolehan Bintang 3 (CUKUP) menduduki posisi ketiga
- Perolehan Bintang 4 (BAGUS) menduduki posisi keempat
- Perolehan Bintang 2 (KURANG) menduduki posisi kelima

Sehingga dari hasil analisis diatas, maka untuk review pengguna Zoom Cloud Meeting dengan menggunakan Algoritma *Naïve Bayes* termasuk kategori “Sangat Bagus”.

4.5.1.4 Hasil Pengujian Zoom Cloud Meeting dengan Model *k-Nearest Neighbor (k-NN)*

Untuk hasil analisis review pengguna Zoom Cloud Meeting dengan menggunakan algoritma *k-Nearest Neighbor (k-NN)* adalah sebagai berikut:

Accuracy : 64,95 %

Kappa : 0,196

Tabel 4.11. Hasil review pengguna Zoom Cloud Meeting dengan Algoritma *k-Nearest Neighbor (k-NN)*

	true 5	true 1	true 2	true 4	true 3	class precision
pred. 5	244	38	8	40	30	67.78%
pred. 1	3	17	3	6	4	51.52%
pred. 2	0	0	0	0	0	0.00%
pred. 4	2	3	0	0	1	0.00%
pred. 3	0	1	0	3	2	33.33%
class recall	97.99%	28.81%	0.00%	0.00%	5.41%	

$$\text{Acc (Accuracy)} = \frac{263}{405} = 0,649$$

Hasil akurasi yang didapatkan sejumlah 64,95 %, dari tabel 4.11 maka hasil class precision dapat disimpulkan sebagai berikut:

- Perolehan Bintang 5 (SANGAT BAGUS) menduduki posisi pertama
- Perolehan Bintang 1 (SANGAT KURANG) menduduki posisi kedua
- Perolehan Bintang 3 (CUKUP) menduduki posisi ketiga
- Perolehan Bintang 4 (BAGUS) menduduki posisi keempat
- Perolehan Bintang 2 (KURANG) menduduki posisi kelima

Sehingga dari hasil analisis diatas, maka untuk review pengguna Zoom Cloud Meeting dengan menggunakan Algoritma Algoritma *k-Nearest Neighbor (k-NN)* termasuk kategori “Sangat Bagus”.

4.5.1.5 Hasil Pengujian Google Meet dengan Model *Naïve Bayes (NB)* dengan *feature selection SMOTE* dan *feature optimasi PSO*

Untuk hasil analisis review pengguna Google Meet dengan menggunakan algoritma *Naïve Bayes* dengan menggunakan *feature selection SMOTE* dan *feature optimasi PSO* adalah sebagai berikut:

Accuracy : 58,87 %

Kappa : 0,440

Tabel 4.12. Hasil review pengguna Google Meet dengan Algoritma *Naïve Bayes+SMOTE+PSO*

	true 1	true 5	true 3	true 4	true 2	class precision
pred. 1	70	42	14	8	0	52.24%
pred. 5	16	100	5	12	0	75.19%
pred. 3	24	29	2	6	0	3.28%
pred. 4	11	52	6	6	0	8.00%
pred. 2	31	31	8	7	254	76.74%
class recall	46.05%	39.37%	5.71%	15.38%	100.00%	
Acc (Accuracy)	$= \frac{432}{734} = 0,588$					

Hasil akurasi yang didapatkan sejumlah 58,87 %, dari tabel 4.12 maka hasil class precision dapat disimpulkan sebagai berikut:

- Perolehan Bintang 2 (KURANG) menduduki posisi pertama
- Perolehan Bintang 5 (SANGAT BAGUS) menduduki posisi kedua
- Perolehan Bintang 1 (SANGAT KURANG) menduduki posisi ketiga
- Perolehan Bintang 4 (BAGUS) menduduki posisi Keempat
- Perolehan Bintang 3 (CUKUP) menduduki posisi kelima

Sehingga dari hasil analisis diatas, maka untuk review pengguna Zoom Cloud Meeting dengan menggunakan Algoritma Algoritma Google Meet dengan Algoritma *Naïve Bayes+SMOTE+PSO* termasuk kategori “Kurang”.

4.5.1.6 Hasil Pengujian Google Meet dengan Model *k-Nearest Neighbor (k-NN)* dengan *feature selection SMOTE* dan *feature optimasi PSO*

Untuk hasil analisis review pengguna Google Meet dengan menggunakan algoritma *k-Nearest Neighbor (k-NN)* dengan menggunakan *feature selection SMOTE* dan *feature optimasi PSO* adalah sebagai berikut:

Accuracy : 73,84%

Kappa : 0,611

Tabel 4.13. Hasil review pengguna Google Meet dengan Algoritma *k-Nearest Neighbor (k-NN)+SMOTE+PSO*

	true 1	true 5	true 3	true 4	true 2	class precision
pred. 1	71	9	7	4	5	73.96%
pred. 5	50	229	14	30	7	69.39%
pred. 3	1	0	0	0	0	0.00%
pred. 4	0	1	0	0	0	0.00%
pred. 2	30	15	14	5	242	79.08%
class recall	46.71%	90.16%	0.00%	0.00%	95.28%	

$$\text{Acc (Accuracy)} = \frac{542}{734} = 0,738$$

Hasil akurasi yang didapatkan sejumlah 73,84 %, dari tabel 4.13 maka hasil class precision dapat disimpulkan sebagai berikut:

- Perolehan Bintang 2 (KURANG) menduduki posisi pertama
- Perolehan Bintang 5 (SANGAT BAGUS) menduduki posisi kedua
- Perolehan Bintang 1 (SANGAT KURANG) menduduki posisi ketiga
- Perolehan Bintang 4 (BAGUS) menduduki posisi Keempat
- Perolehan Bintang 3 (CUKUP) menduduki posisi kelima

Sehingga dari hasil analisis diatas, maka untuk review pengguna Zoom Cloud Meeting dengan menggunakan Algoritma Algoritma Google Meet dengan Algoritma *k-Nearest Neighbor (k-NN)+SMOTE+PSO* termasuk kategori “Kurang”.

4.5.1.7 Hasil Pengujian Zoom Cloud Meeting dengan Model *Naïve Bayes (NB)* dengan *feature selection SMOTE* dan *feature optimasi PSO*

Untuk hasil analisis review pengguna Zoom Cloud Meeting dengan menggunakan algoritma *Naïve Bayes* dengan menggunakan *feature selection SMOTE* dan *feature optimasi PSO* adalah sebagai berikut:

Accuracy : 61,60 %

Kappa : 0,454

Tabel 4.14. Hasil review pengguna Zoom Cloud Meeting dengan Algoritma *Naïve Bayes+SMOTE+PSO*

	true 5	true 1	true 2	true 4	true 3	class precision
pred. 5	104	10	0	13	14	73.76%
pred. 1	18	22	0	7	6	41.51%
pred. 2	45	9	249	7	6	78.80%
pred. 4	53	10	0	13	3	16.46%
pred. 3	29	8	0	9	8	14.81%
class recall	41.77%	37.29%	100.00%	26.53%	21.62%	
Acc (Accuracy)	$= \frac{396}{643} = 0,615$					

Hasil akurasi yang didapatkan sejumlah 61,60 %, dari tabel 4.14 maka hasil class precision dapat disimpulkan sebagai berikut:

- Perolehan Bintang 2 (KURANG) menduduki posisi pertama
- Perolehan Bintang 5 (SANGAT BAGUS) menduduki posisi kedua
- Perolehan Bintang 1 (SANGAT KURANG) menduduki posisi ketiga
- Perolehan Bintang 4 (BAGUS) menduduki posisi Keempat
- Perolehan Bintang 3 (CUKUP) menduduki posisi kelima

Sehingga dari hasil analisis diatas, maka untuk review pengguna Zoom Cloud Meeting dengan menggunakan Algoritma Algoritma Google Meet dengan Algoritma *Naïve Bayes SMOTE+PSO* termasuk kategori “Kurang”.

4.5.1.8 Hasil Pengujian Zoom Cloud Meeting dengan Model *k-Nearest Neighbor (k-NN)* dengan *feature selection SMOTE* dan *feature optimasi PSO*

Untuk hasil analisis review pengguna Zoom Cloud Meeting dengan menggunakan algoritma *k-Nearest Neighbor (k-NN)* dengan menggunakan *feature selection SMOTE* dan *feature optimasi PSO* adalah sebagai berikut:

Accuracy : 77,92%

Kappa : 0,644

Tabel 4.15. Hasil review pengguna Zoom Cloud Meeting dengan Algoritma *k-Nearest Neighbor (k-NN)*

	true 5	true 1	true 2	true 4	true 3	class precision
pred. 5	242	46	1	44	32	66.30%
pred. 1	2	6	0	0	0	75.00%
pred. 2	3	7	248	3	2	94.30%
pred. 4	2	0	0	2	0	50.00%
pred. 3	0	0	0	0	3	100.00%
class recall	97.19%	10.17%	99.60%	4.08%	8.11%	

$$\text{Acc (Accuracy)} = \frac{396}{643} = 0,779$$

Hasil akurasi yang didapatkan sejumlah 61,60 %, dari tabel 4.15 maka hasil class precision dapat disimpulkan sebagai berikut:

- Perolehan Bintang 2 (KURANG) menduduki posisi pertama
- Perolehan Bintang 5 (SANGAT BAGUS) menduduki posisi kedua
- Perolehan Bintang 1 (SANGAT KURANG) menduduki posisi ketiga
- Perolehan Bintang 3 (CUKUP) menduduki posisi Keempat
- Perolehan Bintang 4 (BAGUS) menduduki posisi kelima

Sehingga dari hasil analisis diatas, maka untuk review pengguna Zoom Cloud Meeting dengan menggunakan Algoritma Algoritma Google Meet dengan Algoritma *Naïve Bayes SMOTE+PSO* termasuk kategori “Kurang”.

4.5.1.9 Rangkuman Hasil Pengujian Model Pertama

Dari hasil pengujian review pengguna Google Meet dan Zoom Cloud Meeting dengan algoritma *NB+k-NN* dan *NB+k-NN+SMOTE+PSO*, maka dapat dirangkum hasilnya seperti dibawah ini:

- Hasil Pengujian Google Meet dengan Model *Naïve Bayes (NB)*
 - Accuracy : 34,52 %
 - Kappa : 0,123
 - Hasil : Sangat Bagus
- Hasil Pengujian Google Meet dengan Model *k-Nearest Neighbor (k-NN)*
 - Accuracy : 67,25 %
 - Kappa : 0,425
 - Hasil : Sangat Bagus
- Hasil Pengujian Zoom Cloud Meeting dengan Model *Naïve Bayes (NB)*

- Accuracy* : 33,34 %
Kappa : 0,056
 Hasil : Sangat Bagus
- d. Hasil Pengujian Zoom Cloud Meeting dengan Model *k-Nearest Neighbor (k-NN)*
- Accuracy* : 64,95 %
Kappa : 0,196
 Hasil : Sangat Bagus
- e. Hasil Pengujian Google Meet dengan Model *Naïve Bayes (NB)* dengan *feature selection SMOTE* dan *feature optimasi PSO*
- Accuracy* : 58,87 %
Kappa : 0,440
 Hasil : Kurang
- f. Hasil Pengujian Google Meet dengan Model *k-Nearest Neighbor (k-NN)* dengan *feature selection SMOTE* dan *feature optimasi PSO*
- Accuracy* : 73,84%
Kappa : 0,611
 Hasil : Kurang
- g. Hasil Pengujian Zoom Cloud Meeting dengan Model *Naïve Bayes (NB)* dengan *feature selection SMOTE* dan *feature optimasi PSO*
- Accuracy* : 61,60 %
Kappa : 0,454
 Hasil : Kurang
- h. Hasil Pengujian Zoom Cloud Meeting dengan Model *k-Nearest Neighbor (k-NN)* dengan *feature selection SMOTE* dan *feature optimasi PSO*
- Accuracy* : 77,92%
Kappa : 0,644
 Hasil : Kurang

Jika ditampilkan dalam tabel, maka akan tampak seperti tabel dibawah ini

Tabel 4.16. Hasil Pengujian *Accuracy* dan *Kappa*

Algoritma	Google Meet		Zoom Cloud Meeting	
	Accuracy	Kappa	Accuracy	Kappa
<i>Naïve Bayes</i>	34,52 %	0,123	33,34 %	0,056

Hasil Recall & Preccsian	Sangat Bagus		Sangat Bagus	
<i>Naïve Bayes</i> +SMOTE+PSO	58,87 %	0,440	61,60 %	0,454
Hasil Recall & Preccsian	Kurang		Kurang	
K-Nearest Neighbor	67,25 %	0,425	64,95 %	0,196
Hasil Recall & Preccsian	Sangat Bagus		Sangat Bagus	
K-Nearest Neighbor+SMOTE+PSO	73,84%	0,611	77,92%	0,644
Hasil Recall & Preccsian	Kurang		Kurang	

Dari tabel diatas, dapat dilihat bahwa:

1. Analisis review pengguna Google Meet jika menggunakan algoritma *Naïve Bayes* menunjukkan hasil “sangat bagus”, namun jika menggunakan *feature* SMOTE dan PSO hasilnya menjadi “kurang” sehingga terjadi penurunan;
2. Analisis review pengguna Google Meet jika menggunakan algoritma k-Nearest Neighbor (k-NN) menunjukkan hasil “sangat bagus”, namun jika menggunakan *feature* SMOTE dan PSO hasilnya menjadi “kurang” sehingga terjadi penurunan;
3. Analisis review pengguna Zoom Cloud Meeting jika menggunakan algoritma *Naïve Bayes* menunjukkan hasil “sangat bagus”, namun jika menggunakan *feature* SMOTE dan PSO hasilnya naik menjadi “kurang”;
4. Analisis review pengguna Zoom Cloud Meeting jika menggunakan algoritma k-Nearest Neighbor (k-NN) menunjukkan hasil “sangat bagus”, namun jika menggunakan *feature* SMOTE dan PSO hasilnya naik menjadi “kurang”.

Sehingga dapat ditarik kesimpulan bahwa *feature selection* SMOTE dan *feature optimasi* PSO untuk penggunaan Algoritma *Naïve Bayes* (NB) dan k-Nearest Neighbor (k-NN) pada analisis review pengguna Google Meet menunjukkan penurunan dari “Sangat Bagus” menjadi “Kurang”, sedangkan untuk analisis review pengguna Zoom Cloud Meeting menunjukkan penurunan dari “Sangat bagus” menjadi “kurang”, sehingga penggunaan *feature selection* SMOTE dan *feature optimasi* PSO untuk algoritma *Naïve Bayes* (NB) dan k-Nearest Neighbor (k-NN) sangat berpengaruh terhadap analisis pengguna Google Meet dan Zoom Cloud Meeting.

4.5.2 Hasil Pengujian Metode Kedua – 2 Kelas

4.5.2.1 Hasil Pengujian Google Meet dengan Model *Naïve Bayes*

Untuk hasil analisis review pengguna Google Meet dengan menggunakan algoritma *Naïve Bayes* adalah sebagai berikut:

Accuracy : 66,74 %

AUC : 0,762

Tabel 4.17. Hasil Review Pengguna Google Meet dengan algoritma *Naïve Bayes*

	true 1	true 5	class precision
pred. 1	129	112	53.53%
pred. 5	23	142	86.06%
class recall	84.87%	55.91%	

$$\text{Accuracy} = \frac{271}{406} = 0,667$$

4.5.2.2 Hasil Pengujian Google Meet dengan Model *k-Nearest Neighbor (k-NN)*

Untuk hasil analisis review pengguna Google Meet dengan menggunakan algoritma *k-Nearest Neighbor (k-NN)* adalah sebagai berikut:

Accuracy : 83,24 %

AUC : 0,851

Tabel 4.18. Hasil Review Pengguna Google Meet dengan algoritma *K-Nearest Neighbor*

	true 1	true 5	class precision
pred. 1	113	29	79.58%
pred. 5	39	225	85.23%
class recall	74.34%	88.58%	

$$\text{Accuracy} = \frac{338}{406} = 0,832$$

4.5.2.3 Hasil Pengujian Google Meet dengan Model *Naïve Bayes* dengan *feature selection SMOTE* dan *feature optimasi PSO*

Untuk Hasil analisis review pengguna Google Meet dengan algoritma *Naïve Bayes* dengan *feature selection SMOTE* dan *feature optimasi PSO* adalah sebagai berikut:

Accuracy : 79,33 %

AUC : 0,852

Tabel 4.19. Hasil Review Pengguna Google Meet dengan algoritma *Naïve Bayes* *feature selection SMOTE* dan *feature optimasi PSO*

	true 1	true 5	class precision
pred. 1	236	87	73.07%

pred. 5	18	167	90.27%
class recall	92.91%	65.75%	

$$\text{Accuracy} = \frac{403}{508} = 0,793$$

4.5.2.4 Hasil Pengujian Google Meet dengan Model K-Nearest Neighbor (k-NN) dengan feature selection SMOTE dan feature optimasi PSO

Untuk hasil analisis review pengguna Google Meet menggunakan algoritma K-Nearest Neighbor dengan feature selection SMOTE dan feature optimasi PSO adalah sebagai berikut :

Accuracy : 83,46 %

AUC : 0,911

Tabel 4.20. Hasil Review Pengguna Google Meet dengan algoritma *K-Nearest Neighbor (k-NN) dengan feature selection SMOTE dan feature optimasi PSO*

	true 1	true 5	class precision
pred. 1	195	25	88.64%
pred. 5	59	229	79.51%
class recall	76.77%	90.16%	

$$\text{Accuracy} = \frac{424}{508} = 0,834$$

4.5.2.5 Hasil Pengujian Zoom Cloud Meeting dengan Model Naive Bayes

Untuk hasil analisis review pengguna Zoom Cloud Meeting menggunakan algoritma *Naive Bayes* adalah sebagai berikut:

Accuracy : 72,42 %

AUC : 0,705

Tabel 4.21. Hasil Review Pengguna Zoom Cloud Meeting dengan algoritma *Naive Bayes*

	true 1	true 5	class precision
pred. 1	44	70	38.60%
pred. 5	15	179	92.27%
class recall	74.58%	71.89%	

$$\text{Accuracy} = \frac{223}{308} = 0,724$$

4.5.2.6 Hasil Pengujian Zoom Cloud Meeting dengan Model k-Nearest Neighbor (k-NN)

Untuk hasil analisis review pengguna Zoom Cloud Meeting menggunakan k-nearest neighbor adalah sebagai berikut:

Accuracy : 79,51 %

AUC : 0,696

Tabel 4.22. Hasil Review Pengguna Zoom Cloud Meeting dengan algoritma *k-Nearest Neighbor (k-NN)*

	true 1	true 5	class precision
pred. 1	32	36	47.06%
pred. 5	27	213	88.75%
class recall	54.24%	85.54%	

$$\text{Accuracy} = \frac{245}{308} = 0,795$$

4.5.2.7 Hasil Pengujian Zoom Cloud Meeting dengan Model Naïve Bayes dengan feature selection SMOTE dan feature optimasi PSO

Untuk hasil analisis review pengguna Zoom Cloud Meeting menggunakan Algoritma *Naïve Bayes* dengan feature selection SMOTE dan feature optimasi PSO adalah Sebagai berikut:

Accuracy : 85,76 %

AUC : 0,961

Tabel 4.23. Hasil Review Pengguna Zoom Cloud Meeting dengan algoritma *Naïve Bayes* dengan feature selection SMOTE dan feature optimasi PSO

	true 1	true 5	class precision
pred. 1	248	70	77.99%
pred. 5	1	179	99.44%
class recall	99.60%	71.89%	

$$\text{Accuracy} = \frac{427}{498} = 0,857$$

4.5.2.8 Hasil Pengujian Zoom Cloud Meeting dengan Model *k-Nearest Neighbor (k-NN)* dengan feature selection SMOTE dan feature optimasi PSO

Untuk hasil analisis review pengguna Zoom Cloud meeting menggunakan algoritma *K-Nearest Neighbor* dengan feature selection SMOTE dan feature optimasi PSO adalah sebagai berikut :

Accuracy : 91,56 %

AUC : 0,955

Tabel 4.24. Hasil Review Pengguna Zoom Cloud Meeting dengan algoritma *k-Nearest Neighbor (k-NN)* dengan feature selection SMOTE dan feature optimasi PSO

	true 1	true 5	class precision
pred. 1	228	21	91.57%
pred. 5	21	228	91.57%
class recall	91.57%	91.57%	

$$\text{Accuracy} = \frac{456}{498} = 0,915$$

4.5.3 Hasil Evaluasi Model *Naïve Bayes* (NB) dan *k-Nearest Neighbor* (k-NN)

Berdasarkan hasil pengujian model pertama (5 kelas), maka algoritma *Naïve Bayes* (NB) dan *k-Nearest Neighbor* (k-NN) mendapatkan nilai Accuracy dan Kappa, dimana hasil desain model *Naïve Bayes*, *K-Nearest Neighbor*, *Naïve Bayes+SMOTE+PSO* dan *K-Nearest Neighbor+SMOTE+PSO* mendapatkan hasil seperti pada tabel 4.25 dibawah ini.

Tabel 4.25. Perbandingan nilai Akurasi dan Kappa

Algoritma	Accuracy		Kappa	
	Google Meet	Zoom Cloud Meeting	Google Meet	Zoom Cloud Meeting
<i>Naïve Bayes</i>	34,52 %	33,34 %	0,123	0,056
KNN	64,86 %	61,02 %	0,390	0,200
<i>Naïve Bayes+smote+PSO</i>	58,87 %	61,60 %	0,440	0,454
KNN + Smote+PSO	73,84%	77,92%	0,611	0,644

Sedangkan untuk hasil pengujian model kedua (2 kelas), maka algoritma *Naïve Bayes* (NB) dan *k-Nearest Neighbor* (k-NN) mendapatkan nilai Accuracy dan AUC, dimana hasil desain model *Naïve Bayes*, *K-Nearest Neighbor*, *Naïve Bayes+SMOTE+PSO* dan *K-Nearest Neighbor+SMOTE+PSO* mendapatkan hasil seperti pada tabel 4.26 dibawah ini.

Tabel 4.26. Perbandingan nilai Akurasi dan AUC

Algoritma	Accuracy		AUC	
	Google Meet	Zoom Cloud Meeting	Google Meet	Zoom Cloud Meeting
<i>Naïve Bayes</i>	66,74 %	72,42%	0,762	0,705
KNN	83,24 %	79,51%	0,851	0,696

<i>Naïve Bayes</i> + smote+PSO	79,33%	85,76 %	0,852	0,961
KNN + Smote+ PSO	83,46 %	91,56 %	0,911	0,955

Dapat disimpulkan bahwa feature selection SMOTE dan feature optimasi PSO sangat berpengaruh untuk meningkatkan nilai Accuracy, Kappa maupun AUC, namun sebaliknya jika tanpa menggunakan feature selection dan feature optimasi tidak mempengaruhi pada peningkatan nilai Accuracy, Kappa maupun AUC pada model algoritma *Naïve Bayes* (NB) dan k-Nearest Neighbor (k-NN). Dari hasil pengujian dengan metode yang berbeda (metode 5 kelas dan metode 2 kelas), maka dapat dihasilkan sebagai berikut:

1. Untuk pengujian dengan metode pertama (5 kelas), maka algoritma yang menghasilkan nilai Accuracy dan Kappa cukup baik adalah dengan menggunakan algoritma K-Nearest Neighbor dengan feature selection SMOTE dan feature optimasi PSO untuk review penggunaan Zoom Cloud Meeting dimana untuk nilai Accuracy menghasilkan 77,92 % dan nilai Kappa menghasilkan 0.644 (Strength of Agreement is good) dibandingkan dengan menggunakan algoritma *Naïve Bayes* dengan feature selection Synthetic Minority Over-Sampling Technique (SMOTE) dan feature optimasi Particle Swarm Optimization (PSO). Dimana untuk peningkatan nilai akurasi sebelum dan sesudah menggunakan feature selection SMOTE upsampling dan feature optimasi PSO dapat disampaikan pada tabel 4.27.
2. Untuk pengujian dengan metode kedua (2 kelas), maka algoritma yang menghasilkan nilai Accuracy dan AUC sangat baik adalah dengan menggunakan algoritma K-Nearest Neighbor dengan feature selection SMOTE dan feature optimasi PSO untuk review penggunaan Zoom Cloud Meeting dimana untuk nilai Accuracy menghasilkan 91,56% dan nilai AUC menghasilkan 0,955 dibandingkan dengan menggunakan algoritma *Naïve Bayes* dengan feature selection Synthetic Minority Over-Sampling Technique (SMOTE) dan feature optimasi Particle Swarm Optimization (PSO). Dimana untuk peningkatan nilai akurasi sebelum dan sesudah menggunakan feature selection SMOTE upsampling dan feature optimasi PSO dapat disampaikan pada tabel 4.28

Tabel 4.27. Peningkatan Accuracy sebelum dan sesudah menggunakan SMOTE dan PSO (Metode Pertama – 5 Kelas)

Algoritma	Google Meet	Zoom Cloud Meeting
	Accuracy	
<i>Naïve Bayes</i>	34,52 %	33,34 %
<i>Naïve Bayes</i> +SMOTE+PSO	58,87 %	61,60 %
Progres Kenaikan Accuracy	24,35 %	28,26 %
K-Nearest Neighbor	67,25 %	64,95 %
K-Nearest Neighbor+SMOTE+PSO	73,84%	77,92%
Progres Kenaikan Accuracy	6,59 %	12,97 %

Dari tabel diatas peningkatan nilai akurasi dengan menggunakan algoritma *Naïve Bayes* sebelum dan dan sesudah menggunakan feature selection SMOTE Upsampling dan feature optimasi PSO untuk analisis review pengguna Google Meet mencapai 24,35 %, sedangkan untuk review pengguna Zoom Cloud Meeting kenaikan prosentasenya mencapai 28,26 %. Sedangkan untuk penggunaan algoritma k-Nearest Neighbor pengguna Google Meet kenaikan prosentasenya mencapai 6,59 % , sedangkan untuk pengguna Zoom Cloud Meeting kenaikan prosentasenya mencapai 12,97 % jika menggunakan *feature selection* SMOTE Upsampling dan *feature* optimasi PSO.

Tabel 4.28. Peningkatan Accuracy sebelum dan sesudah menggunakan SMOTE dan PSO (Metode Kedua – 2 Kelas)

Algoritma	Google Meet	Zoom Cloud Meeting
	Accuracy	
<i>Naïve Bayes</i>	66,74 %	72,42 %
<i>Naïve Bayes</i> +SMOTE+PSO	79,33 %	85,76 %
Progres Kenaikan Accuracy	12,59 %	13,34 %
K-Nearest Neighbor	83,24 %	79,51 %
K-Nearest Neighbor+SMOTE+PSO	83,46 %	91,56 %
Progres Kenaikan Accuracy	0,22 %	12,05 %

Dari tabel diatas peningkatan nilai akurasi dengan menggunakan algoritma *Naïve Bayes* sebelum dan dan sesudah menggunakan *feature selection SMOTE Upsampling* dan *feature optimasi PSO* untuk analisis review pengguna Google Meet mencapai 12,59 %, sedangkan untuk review pengguna Zoom Cloud Meeting kenaikan prosentasenya mencapai 13,34 %. Sedangkan untuk penggunaan algoritma *k-Nearest Neighbor* pengguna Google Meet kenaikan prosentasenya mencapai hanya 0.22 %, sedangkan untuk pengguna Zoom Cloud Meeting kenaikan prosentasenya mencapai 12,05 % jika menggunakan *feature selection SMOTE Upsampling* dan *feature optimasi PSO*.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat ditarik beberapa kesimpulan bahwa dari hasil perhitungan algoritma klasifikasi *Naïve Bayes* dan k-Nearest Neighbor (k-NN) dengan feature selection Synthetic Minority Over-Sampling Technique (SMOTE) dan feature optimasi Particle Swarm Optimization (PSO) dalam mengolah data review pengguna Google Meet dan Zoom Cloud Meeting yang diambil dari google play sebanyak 909 data (metode pertama) dan 714 data (metode kedua), menunjukkan bahwa adanya peningkatan nilai Accuracy, Kappa maupun AUC setelah menggunakan feature selection SMOTE Upsampling dan feature optimasi PSO. Hasil kesimpulannya sebagai berikut:

1. Hasil Pengujian metode pertama (5 kelas) menunjukkan tingkat akurasi yang paling baik sebesar 67,25 % dengan algoritma K-Nearest Neighbor untuk pengguna Google Meet, pengguna Zoom Cloud Meeting dengan algoritma yang sama yaitu K-nearest neighbor hanya mendapatkan tingkat akurasi sebesar 64,95 %.
2. Hasil Pengujian metode pertama (2 kelas) menunjukkan tingkat akurasi yang paling baik sebesar 83,24 % dengan algoritma K-Nearest Neighbor untuk pengguna Google Meet, pengguna Zoom Cloud Meeting dengan algoritma yang sama yaitu K-nearest neighbor hanya mendapatkan tingkat akurasi sebesar 79,51 %.
3. Usulan penggunaan feature selection synthetic minority over-sampling technique (SMOTE) dan feature optimasi Particle swarm optimization (PSO) pada algoritma klasifikasi terbukti sangat berpengaruh untuk meningkatkan akurasi pada algoritma *Naïve Bayes* dan k-nearest neighbor untuk pengolahan data review pengguna Google Meet dan Zoom Cloud Meeting pada google play berdasarkan perolehan skor bintang.

5.2 Saran

Saran yang dapat disampaikan berdasarkan hasil penelitian ini adalah sebagai berikut:

1. Kerangka kerja yang diusulkan pada penelitian ini dapat digunakan untuk melakukan penelitian-penelitian *Text Mining* dari berbagai kumpulan dokumen teks berbahasa Indonesia yang tidak terstruktur.
2. Hasil penelitian dapat diterapkan secara konkrit untuk membuat sistem klasifikasi otomatis untuk pesan review produk lainnya.
3. Menggunakan metode pemilihan feature yang lain untuk perbandingan feature yang sesuai dari algoritma klasifikasi seperti *Chi Square*, *Mutual Information*, *Down Sampling*, dan lain-lain.
4. Menambahkan metode klasifikasi lain untuk membandingkan dengan metode dari *Supervised Learning* lainnya dari model yang diusulkan, seperti penggunaan Algoritma *C-45*, *Random Forest* dan lainnya.
5. Penelitian lanjutan dengan mengambil jumlah data yang lebih besar lagi dalam kurun waktu lebih dari 3 bulan pada *Google Play*, sehingga akan mendapatkan nilai akurasi yang lebih baik lagi.

DAFTAR PUSTAKA

- [1] F. Gunawan, M. A. Fauzi, and P. P. Adikara, “Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naive Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile),” *Syst. Inf. Syst. Informatics J.*, vol. 3, no. 2, pp. 1–6, 2017, doi: 10.29080/systemic.v3i2.234.
- [2] A. Salam, J. Zeniarja, and R. S. U. Khasanah, “Analisis Sentimen Data Komentar Sosial Media Facebook Dengan K-Nearest Neighbor (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekpress Indonesia),” *Pros. SINTAK*, pp. 480–486, 2018.
- [3] U. Rofiqoh, R. S. Perdana, and M. A. Fauzi, “Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1725–1732, 2017, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/628>.
- [4] Muljono, D. P. Artanti, A. Syukur, A. Prihandono, and D. R. I. M. Setiadi, “Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes,” pp. 8–9, 2018.
- [5] A. Darmawan and S. Syamsiah, “Analisa Model Support Vector Machine Textmining pada Komentar Positif dan Negatif untuk Review PerbandinganWhatsapp Vs BBM,” *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 1, no. 1, pp. 74–82, 2016, doi: 10.30998/string.v1i1.971.
- [6] S. Hanggara, T. M. Akhriza, and M. Husni, “APLIKASI WEB UNTUK ANALISIS SENTIMEN PADA OPINI PRODUK DENGAN METODE NAIVE BAYES CLASSIFIER,” *Semin. Nas. Inov. Dan Apl. Teknol. Di Ind.*, vol. 4, no. 2, p. A33.1-A33.6, 2017.
- [7] W. E. Nurjanah, R. S. Perdana, and M. A. Fauzi, “Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1750–1757, 2017, doi: 10.1074/jbc.M209498200.
- [8] A. Sari, F. V., & Wibowo, “Analisis Sentimen Pelanggan Toko Online Jd. Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi,” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 2, no. 2, pp. 681–686, 2019.
- [9] G. A. Buntoro, “Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter,” *Integer J. Maret*, vol. 1, no. 1, pp. 32–41, 2017, [Online]. Available: https://www.researchgate.net/profile/Ghulam_Buntoro/publication/316617194_Analisis_Sentimen_Calon_Gubernur_DKI_Jakarta_2017_Di_Twitter/links/5907eee44585152d2e9ff992/Analisis-Sentimen-Calon-Gubernur-DKI-Jakarta-2017-Di-Twitter.pdf.
- [10] S. Fanissa, M. A. Fauzi, and S. Adinugroho, “Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 8, pp. 2766–2770, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/1962>.
- [11] M. W. Pertiwi, “Analisis sentimen opini publik mengenai sarana dan

- transportasi mudik tahun 2019 pada twitter menggunakan algoritma naïve bayes, neural network, KNN dan SVM,” *Inti Nusa Mandiri*, vol. 14, no. 1, pp. 27–32, 2019.
- [12] A. M. F. Hulu and K. M. Lhaksana, “Analisis Sentimen pada Twitter Menggunakan Metode Support Vector Machine,” *Semin. Nas. Teknol. Fak. Tek. Univ. Krisnadwipayana*, vol. 1, no. 1, pp. 739–742, 2019.
- [13] G. Gunadi and D. I. Sensuse, “Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) : STUDI KASUS PERCETAKAN PT. GRAMEDIA,” *Telematika*, vol. 4, no. 1, pp. 118–132, 2012.
- [14] Maharani *et al.*, “IMPLEMENTASI DATA MINING UNTUK PENGATURAN LAYOUT MINIMARKET DENGAN MENERAPKAN ASSOCIATION RULE,” vol. 4, no. 4, pp. 6–11, 2017.
- [15] Y. Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4.5,” *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017.
- [16] G. Abdillah, F. A. Putra, and F. Renaldi, “Penerapan Data Mining Pemakaian Air Pelanggan Untuk Menentukan Klasifikasi Potensi Pemakaian Air Pelanggan Baru Di Pdam Tirta Raharja Menggunakan Algoritma K-Means,” *Sentika 2016*, vol. 2016, no. Sentika, pp. 18–19, 2016.
- [17] Hartanto, “Text Mining Dan Sentimen Analisis Twitter Pada Gerakan Lgbt,” *Intuisi J. Psikol. Ilm.*, vol. 9, no. 1, pp. 18–25, 2017.
- [18] N. L. Ratniasih, M. Sudarma, and N. Gunantara, “Penerapan Text Mining Dalam Spam Filtering Untuk Aplikasi Chat,” *Maj. Ilm. Teknol. Elektro*, vol. 16, no. 3, p. 13, 2017, doi: 10.24843/mite.2017.v16i03p03.
- [19] I. A. Marie, L. Hakim, D. Sugiarto, and W. Septiani, “Perbandingan Performansi Teknik Klasifikasi Breakdown Mesin pada Proses Produksi Pembuatan Battery Mobil,” *J. Ilm. Tek. Ind.*, vol. 18, no. 1, pp. 33–41, 2019, doi: 10.23917/jiti.v18i1.7232.
- [20] M. M. Saritas and A. Yasar, “Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification,” *Faraday Discuss.*, vol. 2, no. 7, pp. 88–91, 2019, doi: 10.1039/b000000x.
- [21] Sulastri and Y. S. Nugroho, “Penerapan Data Mining Untuk Prediksi Rating Penjualan Buku Menggunakan Metode Naive Bayes,” *Duta.com*, vol. 12, no. 2, pp. 57–72, 2017.
- [22] Minarni and P. Irawan, “Implementasi Metode Naive Bayes Untuk Diagnosa Penyakit Lambung,” *J. TEKNOIF*, vol. 9986, no. September, pp. 115–123, 2019, doi: 10.21063/JTIF.2019.V7.2.
- [23] J. Riany, M. Fajar, and M. P. Lukman, “Penerapan Deep Sentiment Analysis pada Angket Penilaian Terbuka Menggunakan K-Nearest Neighbor,” *Sisfo*, vol. 06, no. 01, pp. 147–156, 2016, doi: 10.24089/j.sisfo.2016.09.011.
- [24] E. Hardiyanto and F. Rahutomo, “STUDI AWAL KLASIFIKASI ARTIKEL WIKIPEDIA BAHASA INDONESIA DENGAN MENGGUNAKAN METODA K NEAREST NEIGHBOR,” vol. 27, no. 1, pp. 15–16, 2016.
- [25] N. H. A. Sari, M. A. F. Fauzi, and P. P. Adikara, “Klasifikasi Dokumen Sambat Online Menggunakan Metode K-Nearest Klasifikasi Dokumen Sambat Online Menggunakan Metode K-Nearest Neighbor dan Features Selection Berbasis Categorical Proportional Difference,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. August, pp. 2449–2454, 2018.

- [26] A. Zanuardi and H. Suprayitno, "Analisa Karakteristik Kecelakaan Lalu Lintas di Jalan Ahmad Yani Surabaya melalui Pendekatan Knowledge Discovery in Database," *J. Manajemen Aset Infrastruktur Fasilitas*, vol. 2, no. 1, pp. 45–55, 2018, doi: 10.12962/j26151847.v2i1.3767.
- [27] L. Ardiantoro, S. Zahara, and N. Sunarmi, "Pemanfaatan Knowledge Data Discovery(KDD) Pada Pola Permainan Atlet Bulutangkis," *Explor. IT J. Keilmuan dan Apl. Tek. Inform.*, vol. 11, no. 1, pp. 1–6, 2019, doi: 10.35891/explorit.v11i1.1467.
- [28] A. Putra, "SOLUSI PREDIKSI MAHASISWA DROP OUT PADA PROGRAM STUDI SISTEM INFORMASI FAKULTAS ILMU KOMPUTER UNIVERSITAS BINA DARMA," *SIMETRIS*, vol. 8, no. 3, pp. 177–184, 2017, doi: 10.1103/PhysRevE.94.033206.
- [29] S. Haryati, A. Sudarsono, and E. Suryana, "IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI MASA STUDI MAHASISWA MENGGUNAKAN ALGORITMA C4.5 (STUDI KASUS: UNIVERSITAS DEHASSEN BENGKULU)," *J. Media Infotama*, vol. 11, no. 2, pp. 130–138, 2015.
- [30] N. P. Wong, F. N. S. Damanik, E. S. Jaya, and R. Rajaya, "Perbandingan Algoritma C4 . 5 dan Classification and Regression Tree (CART) Dalam Menyeleksi Calon Karyawan," vol. 20, no. 1, pp. 11–18, 2019.
- [31] N. Abdulloh and A. F. Hidayatullah, "Deteksi Cyberbullying pada Cuitan Media Sosial Twitter," vol. Vol 1, No, 2019.
- [32] M. F. Mujaddid, Adiwijaya, and S. Al-faraby, "ANALISIS CHURN PREDICTION MENGGUNAKAN METODE LOGISTIC REGRESSION DAN SMOTE (Synthetic Minority Over-sampling Technique) PADA PERUSAHAAN TELEKOMUNIKASI," vol. 4, no. 3, pp. 5046–5054, 2017.
- [33] J. Mathew, M. Luo, C. K. Pang, and H. L. Chan, "Kernel-based SMOTE for SVM classification of imbalanced datasets," *IECON 2015 - 41st Annu. Conf. IEEE Ind. Electron. Soc.*, pp. 1127–1132, 2015, doi: 10.1109/IECON.2015.7392251.
- [34] H. Muhamad, C. A. Prasojo, N. A. Sugianto, L. Surtiningsih, and I. Cholissodin, "Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 3, p. 180, 2017, doi: 10.25126/jtiik.201743251.
- [35] R. F. Margeritha, R. S. Hartati, and N. P. Satriya Utama, "Analisis Penyambungan Distributed Generation Guna Meminimalkan Rugi-Rugi Daya Menggunakan Metode Particle Swarm Optimization (PSO)," *Maj. Ilm. Teknol. Elektro*, vol. 16, no. 3, p. 122, 2017, doi: 10.24843/mite.2017.v16i03p19.
- [36] Y. E. Achyani, "Penerapan Metode Particle Swarm Optimization Pada Optimasi Prediksi Pemasaran Langsung," *J. Inform.*, vol. 5, no. 1, pp. 1–11, 2018, doi: 10.31311/ji.v5i1.2736.
- [37] B. Rahmat C.T.I. *et al.*, "Implemetasi k-means clustering pada rapidminer untuk analisis daerah rawan kecelakaan," *Semin. Nas. Ris. Kuantitatif Terap. 2017*, no. April, pp. 58–60, 2017.
- [38] J. R. Raco, "Metode Penelitian Kualitatif: Jenis, Karakteristik dan Keunggulannya," 2010.

DAFTAR RIWAYAT HIDUP

A. Biodata Mahasiswa

N.I.M : 114002242
 Nama Lengkap : MUHAMMAD REZKI
 Tempat dan Tanggal Lahir : MAKRAMPAI, 15 SEPTEMBER 1995
 Alamat Lengkap : Dusun Melur Rt.005/03 Mekar Sekuntum
 Tebas Kabupaten Sambas

B. Riwayat Pendidikan Formal dan Non Formal

1. SDN 39 Sejiram, lulus tahun 2008
2. SMPN 1 Tebas, lulus tahun 2011
3. SMKN 1 Tebas, lulus tahun 2014
4. D3 BSI Pontianak, lulus tahun 2017
5. S1 BSI BANDUNG, lulus tahun 2018

C. Riwayat Berorganisasi / Pekerjaan

1. Anggota Himmi Amik BSI Pontianak
2. Magang BCC Pusat
3. Staf kemahasiswaan Bsi Pusat
4. Staf BIC Pusat

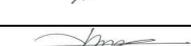
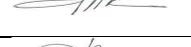


Jakarta, 04 Agustus 2020

Muhammad Rezki

	LEMBAR BIMBINGAN TESIS
	STMIK NUSA MANDIRI

NIM : 14002242
 Nama Lengkap : Muhammad Rezki
Dosen Pembimbing I : Dr.Tr. Lindung Parningotan Manik, M.T.I.
 Judul Tugas Akhir : Analisis Review Pengguna Google Meet Dan Zoom
 Cloud Meeting Menggunakan Algoritma Naïve Bayes
 Dan K-Nearest Neighbor

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing
1.	06 April 2020	Bimbingan perdana	
2.	30 April 2020	Revisi judul	
3.	25 Mei 2020	Pengajuan bab I	
4.	16 Juli 2020	Acc Bab I & Pengajuan Bab II	
5.	20 Juli 2020	Acc Bab II & Pengajuan Bab III	
6.	21 Juli 2020	Acc Bab III & Pengajuan Bab IV	
7.	23 Juli 2020	Acc Bab IV & Pengajuan Bab V	
8.	06 Agustus 2020	Bimbingan Terakhir & Acc Keseluruhan	

Catatan untuk Dosen Pembimbing.

Bimbingan Tugas Akhir

- Dimulai pada tanggal : 06 April 2020
- Diakhiri pada tanggal : 06 Agustus 2020
- Jumlah pertemuan bimbingan : 8 Kali Pertemuan

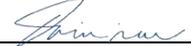
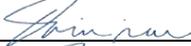
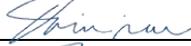
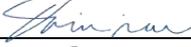
Disetujui oleh,
Dosen Pembimbing I



(Dr.Tr. Lindung Parningotan Manik, M.T.I.)

	LEMBAR BIMBINGAN TESIS
	STMIK NUSA MANDIRI

NIM : 14002242
 Nama Lengkap : Muhammad Rezki
 Dosen Pembimbing II : Dr. Dwiza Riana, S.Si, MM, M.Kom
 Judul Tugas Akhir : Analisis Review Pengguna Google Meet Dan Zoom
 Cloud Meeting Menggunakan Algoritma Naïve Bayes
 Dan K-Nearest Neighbor

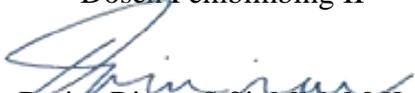
No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing
9.	06 April 2020	Bimbingan perdana	
10.	30 April 2020	Revisi judul	
11.	25 Mei 2020	Pengajuan bab I	
12.	16 Juli 2020	Acc Bab I & Pengajuan Bab II	
13.	20 Juli 2020	Acc Bab II & Pengajuan Bab III	
14.	21 Juli 2020	Acc Bab III & Pengajuan Bab IV	
15.	23 Juli 2020	Acc Bab IV & Pengajuan Bab V	
16.	06 Agustus 2020	Bimbingan Terakhir & Acc Keseluruhan	

Catatan untuk Dosen Pembimbing.

Bimbingan Tugas Akhir

- Dimulai pada tanggal : 06 April 2020
- Diakhiri pada tanggal : 06 Agustus 2020
- Jumlah pertemuan bimbingan : 8 Kali Pertemuan

Disetujui oleh,
Dosen Pembimbing II


 (Dr. Dwiza Riana, S.Si, MM, M.Kom)

Lampiran 1. Dataset Google Meet

no	ulasan	rating
1	Maaf ini ada apa ya kenapa saya nginstal aplikasi tidak.terunduh" padahal kan saya udah pakai akunya... Apa ada yang salah?	1
2	Aplikasi Google meet ini sangat bermanfaat dan membantu dalam memberikan materi kuliah, bimbingan tugas akhir mahasiswa dan membuka ruang konsultasi bagi Dosen pembimbing di kampus. apalagi kita di indonesia sedang dalam situasi Darurat bencana kesehatan masyarakat (dampak pandemi Covid 19) diberlakukan physical distance di Wilayah	5
3	Saya sangat tertarik dengan aplikasi Google meet ini apalagi disaat sekarang ini,ilmu saya masih minim sekali tentang ini , mudah-mudahan kita bisa membawa anak didik kita dengan pembelajarn online ini ,walaupun kita tidak dapat bertatap muka langsung dengan siswa , tapi pembelajaran tetap berlangsung dengan baik, juga bisa memanfaatkan teknologi,	5
4	kalau bisa dikasih fitur unduh untuk setiap file yang dikirim bapak ibu guru	5
5	Jadi saya mencoba membuka kelas saya, tetapi dikatakan bukan kelas saya dan kemudian segera menghapusnya dari pandangan saya. Saya memeriksa ruang kelas di komputer saya untuk melihat apakah ada sesuatu yang salah dengan akun saya, tetapi semua kelas saya ada di sana, jadi saya cukup yakin bahwa aplikasi tidak berfungsi. Ini hanya ketidaknyamanan utama, karena saya tidak dapat menggunakan ponsel saya untuk mentidakses hal-hal seperti pekerjaan rumah dan informasi yang diperlukan. Ini lebih buruk karena saya tidak selalu memiliki akses ke komputer saya.	1
6	Susah gabung	3
7	Aplikasi sangat bermanfaat, buat proses belajar mengajar jadi gampang dan seru	5
8	Bqgus, tapi lambat mendownload	3
9	Aplikasinya sangat gila saya lieur euy	4
10	Tryout online dan mengerjakan PR jadi lebih mudah	5
11	Aplikasi error terus, tidak bisa masuk	1
12	Mohon di tingkatkan aplikasi untuk penulisan rumus kimia dan matematika	4
13	Ngbug,keluar sendiri dari kelas,tiba tiba logout sendiri	1
14	Setelah guru memberikan soal saya malah tidak bisa masuk aneh woi	1
15	Lebih baik google meet menyediakan fitur sticker supaya lebih menarik	5
16	terima kasih sudah di perbaiki permasalahannya	5
17	Lumayan bagus untuk anak 2 yang malas belajar	3
18	aplikasi ini tidak bisa di unduh	5
19	tidak bisa masuk, logo aplikasi nya jadi ada lambang ikan nya	1
20	Kenapa tiba-tiba data kelas di saya ke hapuslah payah	1
21	Cocok buat mahasiswa, sekalipun hapenya kelas kentang, tetep berjalan kencang	5
22	yang bilang aplikasi ini jelek, berarti hp nya kentang	5
23	saya kasih bintang satu soalnya pb saya masih mayor	1
24	Kenapa saya tidak bisa masuk ke akun saya yang lama, tapi kalau lewat google bisa	1
25	Mantap bisa ulangan dirumah	5
26	keluar sendiri, keluar dari kelas	1
27	Ngebug keluar sendiri dari kelas, dan setelah keluar tidak bisa log in tolong diperbaiki Dan kenapa ada logo ikannya? Apa guna lambang ikan itu!!!	1
28	Tiba" keluar dari kelas, penambahan kelas tidak nampak pada aplikasi tetapi saat buka d web bisa.. Tidak bisa dibuka setelah pembaruan	1
29	Google meet nya ngebug, tidak bisa dibuka.Tolong perbaiki bugnya secepat mungkin, saya banyak tugas yang harus di kumpul.	1
30	Ini kenapa pas mau melihat tugas sesuai kelas malah daftar kelas yang ada terhapus ya? terus mau masuk malah tidak bisa lagi pakai akun lain juga tidak bisa.	1
31	Aplikasi habis di update kenapa ngebug, keluar dari kelas sendiri	1

Lampiran 2. Dataset Zoom Cloud Meeting

no	text	rating
1	saya suka sekali, karena sangat membantu meetings jarak jauh, apalagi disaat korona melanda di Indonesia, semua orang tidak boleh keluar rumah.	5
2	Bisa belajar bareng	5
3	Saya baru akan mencoba untuk belajar pakai zoom us Terima kasih	5
4	tidak bisa masuk app nya	1
5	terima kasih ZOOM	5
6	di downlo d nya lama banget ya dari tadi	2
7	bagus mas lur tapi belum banyak teman kantor yang gabung disini	5
8	mahal amat pakai aplikasi nya	1
9	lumayan guys, jadi lebih hemat biaya, waktu dan tenaga. lebih praktis, cuma tidak bisa makan bareng bareng dalam suasana yg sama	4
10	Manfaat, bisa ketemu keluarga	4
11	saya sudah instal terus cara buka untuk memulai di mana ya memulainya	4
12	aplikasinya mantap	5
13	aplikasi ini bisa buat laptop kan? tapi kenapa saya ga bisa? sudah terinstal tapi ga mau kebuka?? kenapa?? dan kasih tau caranya.. Tolong	2
14	layak sekali untuk kegiatan pembelajaran dan rakor hemat biaya dan efisien tepat guna dan tepat waktu, sehingga cocok diterapkasi an dalam keadaan darurat	5
15	coba dulu	3
16	assalamualaikum selamat pagi	5
17	suka	5
18	mantap betul	4
19	belum faham	4
20	sejauh ini bagus	5
21	bagus	5
22	sangat bagus	5
23	download lancar	5
24	Sangat Luar Biasa	5
25	Aku bisa belajar presentasi dari seminar 3i network kereeeeeen	5
26	tes dulu	5
27	Lumayan....	5
28	Sangat membantu	5
29	Megah	5
30	Sangat bermanfaat. Good..!	5
31	Sangat menyenangkan	5
32	mantap aplikasi nya	5
33	Membantu menyelesaikan meeting dg efisien	4
34	Bagus untuk meeting teleconference	5
35	3 dulu.. kalo bagus mau nambahin bintang nya!😊	3
36	Sangat membantu sekali utk training oriflame	5
37	Aplikasi ini enak untuk mengetahui bagaimana cara berbisnis online	5
38	Good for video conference	5
39	belum tahu	5
40	Zoom cloud meetings juuuuooooos banget	5
41	luar biasa	5
42	Excellent	5
43	Log in rumit	1
44	Apl ini sangat membantu	5
45	Memungkinkan seminar melalui online	5
46	okay	5
47	jelek	1
48	Bagus aplikasinya	5