IMPLEMENTASI SENTIMENTAL ANALYSIS DAN SOCIAL NETWORK ANALYSIS PADA OPINI PUBLIK PENGGUNA TWITTER TERHADAP ISU POLITIK BERBASIS WEB



SKRIPSI

SYAIR MUHARRAM

42619022

PROGRAM STUDI D-4 TEKNIK MULTIMEDIA DAN JARINGAN

JURUSAN TEKNIK ELEKTRO

POLITEKNIK NEGERI UJUNG PANDANG

2023

# HALAMAN PERSETUJUAN

Proposal skripsi ini dengan judul Implementasi Sentimental Analysis Dan Social Network Analysis Pada Opini Publik Terhadap Isu Politik Berbasis Web oleh SYAIR MUHARRAM NIM 42619022 dinyatakan layak untuk diseminarkan.

Makassar, 18 September 2023

Pembimbing I, Pembimbing II,

Asriyadi, S.ST.,M.Eng Muh. Ahyar, SST.,MT

NIP. 197708142003121002 NIP. 198410272008121003

Mengetahui

Ketua Program Studi,

Asriyadi, S.ST.,M.Eng

NIP. 197708142003121002

# HALAMAN PENERIMAAN

Pada hari ini, hari Senin tanggal 13 Februari 2023, Tim Penguji Seminar Proposal Skripsi telah menerima dengan baik hasil seminar proposal skripsi oleh mahasiswa: Syair Muharram NIM 42619022 dengan judul Implementasi Sentimental Analysis Dan Network Analysis Opini Pengguna Twitter Pada Isu Politik Berbasis Web.

Makassar, Februari 2023

Tim Seminar Proposal Skripsi:

1. Dr. Ir. Hafsah Nirwana, M.T. Ketua ( )

2. Mardawia M. Parenreng, S.ST., M.T. Sekretaris ( )

3. Ainun Jariyah, S.T., M.T. Anggota ( )

4. Alvian Bastian Anggota ( )

5. Asriyadi, S.ST., M.Eng. Pembimbing I ( )

6. Muh. Ahyar, S.ST., M.T. Pembimbing II ( )

# KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Allah swt. Yang telah melimpahkan rahmat dan karunia-Nya sehingga penulisan proposal skripsi program diploma empat (D-4) bidang Teknik Multimedia dan Jaringan dapat kami selesaikan.

Penyusunan proposal skripsi ini dimulai dengan beberapa kali perbaikan, pemutakhiran hingga proses konsultasi langsung antara penyusun dengan dosen pengarah program studi berdasarkan pedoman dan kaidah yang berlaku di Politeknik Negeri Ujung Pandang, serta referensi terkait. Masukan dan saran yang terkait dengan prosesi penelitian dan skripsi dihimpun sedemikian rupa agar dapat diimplementasikan dan diamalkan oleh penyusun.

Kami berharap bahwa penulisan proposal skripsi ini dapat menjadi acuan dan referensi prosesi penelitian lebih lanjut bagi masyarakat, civitas akademika khususnya penyusun itu sendiri. Akhirnya, kami ingin menyampaikan ucapan terima kasih kepada seluruh pihak yang telah membantu dan memberi fasilitas mulai dari tahap perencanaan hingga penyusunan proposal skripsi ini.

Makassar, Januari 2023

Penulis

# DAFTAR TABEL

Tabel 3.1 Time Schedule 16

# DAFTAR GAMBAR

Gambar 2.1 Arsitektur LSTM 5

Gambar 2.2 Contoh diagram pada *Sentimental Analysis* 6

Gambar 2.3 Contoh implementasi Social Network Analysis pada penyebaran informasi covid-19 8

Gambar 3.1 Metode waterfall 17

Gambar 3.2 Tampilan flowchart sentiment analysis 18

Gambar 3.3 Flowchart Network Analysis 19

Gambar 3.4 Use case diagram 20

Gambar 3.5 Tampilan halaman home 20

Gambar 3.6 Tampilan halaman Analisa sentimen 21

Gambar 3.7 Tampilan halaman visualisasi jika card lihat ditekan 21

Gambar 3.8 Tampilan halaman Network Analysis 22

Gambar 3.9 Tampilan halaman jika card pada Network Analysis ditekan 22

Gambar 3.10 Tampilan halaman admin 23

# DAFTAR ISI

[HALAMAN PERSETUJUAN ii](#_Toc146108289)

[HALAMAN PENERIMAAN iii](#_Toc146108290)

[KATA PENGANTAR iv](#_Toc146108291)

[DAFTAR TABEL v](#_Toc146108292)

[DAFTAR GAMBAR vi](#_Toc146108293)

[DAFTAR ISI vii](#_Toc146108294)

[BAB I PENDAHULUAN 1](#_Toc146108295)

[1.1 Latar Belakang 1](#_Toc146108296)

[1.2 Rumusan Masalah 2](#_Toc146108297)

[1.3 Ruang Lingkup Penelitian 3](#_Toc146108298)

[1.4 Tujuan Penelitian 3](#_Toc146108299)

[1.5 Manfaat Penelitian 3](#_Toc146108300)

[BAB II TINJAUAN PUSTAKA 4](#_Toc146108301)

[2.1 Penelitian Terdahulu 4](#_Toc146108302)

[2.2 Natural Language Processing 4](#_Toc146108303)

[2.3 LSTM 5](#_Toc146108304)

[2.4 Sentimental Analysis 6](#_Toc146108305)

[2.5 Social Network Analysis 7](#_Toc146108306)

[2.6 Twitter 8](#_Toc146108307)

[2.7 Text Mining 9](#_Toc146108308)

[2.8 Python 11](#_Toc146108309)

[2.9 NLTK 12](#_Toc146108310)

[2.10 TwitterWebScraper 13](#_Toc146108311)

[2.11 NetworkX 13](#_Toc146108312)

[2.12 Visual Studio Code 13](#_Toc146108313)

[2.13 Flask 13](#_Toc146108314)

[2.14 MySQL 14](#_Toc146108315)

[2.15 Metode Pengembangan Waterfall 14](#_Toc146108316)

[2.16 Pengujian Black Box 15](#_Toc146108317)

[BAB III METODE PENELITIAN 16](#_Toc146108318)

[3.1 Tempat dan Waktu Penelitian 16](#_Toc146108319)

[3.2 Alat dan Bahan 16](#_Toc146108320)

[3.3 Prosedur Penelitian 17](#_Toc146108321)

[3.3.1 Metode Penelitian dan Tahapan Pengembangan Aplikasi 17](#_Toc146108322)

[3.3.2 Analisis Kebutuhan 17](#_Toc146108323)

[3.3.3 Desain Sistem 18](#_Toc146108324)

[3.3.4 Implementasi dan Pengkodean 23](#_Toc146108325)

[3.3.5 Testing 23](#_Toc146108326)

[3.3.6 Pemeliharaan 23](#_Toc146108327)

[3.4 Pengujian Sistem 23](#_Toc146108328)

[DAFTAR PUSTAKA 25](#_Toc146108329)

# PENDAHULUAN

## Latar Belakang

Majunya perkembangan teknologi informasi saat ini membawa banyak keuntungan dan manfaat bagi masyarakat. Keuntungan dan manfaat yang didapatkan antara lain memudahkan masyarakat untuk berinteraksi dengan banyak orang, memperluas relasi, jarak dan waktu bukan lagi masalah, lebih mudah dalam mengekspresikan diri, dan penyebaran informasi dapat berlangsung secara cepat.

Pada Januari 2022, jumlah pengguna media sosial di Indonesia mencapai 191,4 juta. Jumlah pengguna jejaring sosial di Indonesia pada awal tahun 2022 setara dengan 68,9% dari total penduduk di Indonesia, jumlah pengguna jejaring sosial di Indonesia meningkat sebanyak 21 juta atau 12,5% dari tahun 2021 hingga tahun 2022. (Azhar, 2022).

Pengguna sosial media dapat mengekspresikan opininya pada berbagai bidang isu, misalnya pada isu politik atau isu – isu lainnya pada platform medi sosial. Adapun Media sosial yang dipakai untuk dijadikan wadah beropini sesorang yaitu X atau yang sebelumnya Bernama Twitter. X atau yang selanjutnya akan ditulis sebagai Twitter memiliki struktur postingan yang cukup memudahkan seseorang dapat memberikan opininya. Selain itu pada sisi pihak pengembang aplikasi twitter menawarkan layanan API (*Application Programming Interface*) yang memungkinkan para pengembang aplikasi untuk mendapatkan data twitter secara langsung untuk pengolahan data lebih lanjut. Hal tersebut membuat twitter menjadi salah satu platform yang kerap digunakan untuk proses data mining, salah satu penerapan *data mining* yaitu *Sentimental analysis* dan *Social Network analysis*.

*Sentimental analysis* merupakan proses untuk mengetahui suatu kalimat mengandung tendensi positif, negatif atau netral. Positif berarti adanya aspek suka, setuju, dan sejenisnya terhadap suatu topik (Handayanto & Herlawati, 2020). Sentimental analisis merupakan salah satu cabang dari NLP (*Natural Language Processing*) yang berfungsi untuk mengklasifikan emosi atau sentimen orang – orang berdasarkan data teks yang didapat dari *tweet*, ulasan produk, ataupun komentar. *Sentimental Analysis* digunakan untuk mengetahui seberapa cenderung opini seseorang terhadap suatu isu atau permasalahan, dengan titik kecenderungan berupa, positif, negatif atau netral.

*Social Network Analysis* adalah analisis yang digunakan untuk merepresentasikan hubungan antar beberapa orang, komunitas, atau perusahaan dengan menggunakan teknik analisis grafik (Selisker, 2017). Penerapan *Social Network Analysis* dalam suatu aplikasi yang mampu menggambarkan relasi atau hubungan antar individu dengan melakukan visualisasi dalam bentuk graf kiranya dapat membantu proses pemecahan masalah yang ada. Selain itu, akan dilakukan proses kalkulasi terhadap setiap relasi antar individu untuk menemukan *centrality* dari sebuah jejaring sosial yang didasarkan pada posisi masing-masing individu yang terkait dalam struktur jaringan. Kedua analisis ini dapat digabungkan untuk mengevaluasi opini publik dari topik tertentu dan bagaimana itu terhubung dengan interaksi sosial. Namun sayangnya, untuk saat ini masyarakat kekurangan aplikasi untuk membantu dalam hal yang telah disebutkan diatas.

Oleh karena peneliti memilih judul implementasi sentimental analysis dan social network analysis pada opini publik pengguna twitter terhadap isu politik berbasis web

## Rumusan Masalah

Berdasarkan latar belakang di atas , dapat dirumuskan permasalahan dalam penelitian ini sebagai berikut:

1. Bagaimana implementasi *sentimental analysis* dan *social network analysis pada data tweet* terhadap isu politik
2. Bagaimana visualisasi data *sentimental analysis* dan *social network analysis*
3. Bagaimana proses pengambilan data twitter

## Ruang Lingkup Penelitian

Untuk pembahasan topik yang lebih terarah pada tujuan yang diingin pada penelitian ini, maka penelitian ini diberikan Batasan sebagai berikut:

1. Data *tweet* yang digunakan yakni *tweet* yang berbahasa Indonesia.
2. Opini publik pada data tweet yakni berkaitan dengan isu politik
3. Metode yang digunakan dalam sentimental analysis yaitu metode LSTM
4. Jumlah data tweet yang dipakai sebanyak 1900
5. Menampilkan hasil analisa sentimen dan analisa jaringan dalam bentuk gambar

## Tujuan Penelitian

Adapun tujuan penelitian yang dicapai penulis antara lain:

1. Untuk mengimplementasikan *sentimental analysis* dan *network analysis* pada da*ta twitter*

2. Untuk memvisualisasikan data *sentimental analysis* dan *network analysis*

## Manfaat Penelitian

Penelitian ini diharapkan dapat bermanfaat bagi penulis maupun pembaca tentang gambaran sentimental analysis dan network analysis pada data twitter. selain itu penelitian ini bermanfaat untuk masyarakat yang ingin mengetahui Analisa sentiment dan Analisa jejaring social terkait isu-isu tertentu.

# **TINJAUAN PUSTAKA**

## Penelitian Terdahulu

Primandani Arsi dan Retno Waluyo dalam jurnalnya yang berjudul Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma *Support Vector Machine* (SVM). Dalam penelitian tersebut diusulkan metode Sup*port Vector Machine (*SVM) untuk diterapkan pada *tweets* topik pemindahan ibu kota Indonesia untuk tujuan klasifikasi kelas sentimen pada media sosial Twitter. Berdasarkan hasil dari proses pengujian yang telah dilakukan terhadap *tweets* sentimen pemindahan ibu kota dari media sosial twitter sebanyak 1.236 *tweets* (404 positif dan 832 negatif) menggunakan SVM maka hasil hipotesa awal sesuai hasil akhir yakni SVM lebih baik dari metode sebelumnya (BM25 + KNN dan Naive Bayes).

Jurnal dengan judul Implementasi Social Network Analysis Dalam Penyebaran Informasi Virus Corona (Covid-19) Di Twitter yang ditulis oleh Dwi Inayah dan Fredy Law Purba. Metode yang digunakan adalah *Wordcloud Analysis*, *Sentiment Analysis*, dan *Social Network Analysis*. Hasil *wordcloud analysis* menunjukkan bahwa kata yang sering muncul terkait Covid-19 adalah “positif”, “pandemi” dan “Indonesia”. Hasil *sentiment analysis* menunjukkan bahwa cuitan bersentimen netral merupakan yang terbanyak, disusul cuitan bersentimen negatif, kemudian bersentimen positif.

Nico Munasatya dan Sendi Novianto dalam jurnal dengan judul Natural Language Processing untuk Analisis Sentimen Presiden Jokowi Menggunakan Multi-Layer Perceptron. Hasil Pengujian pada jurnal tersebut menunjukan bahwa akurasi algoritma *multi-layer perceptro*n dengan jumlah *Hidden layer*, *Learning Rate, Shuffling* dan *Drop O*ut berpengaruh dari nilai akurasi yang dihasilkan.

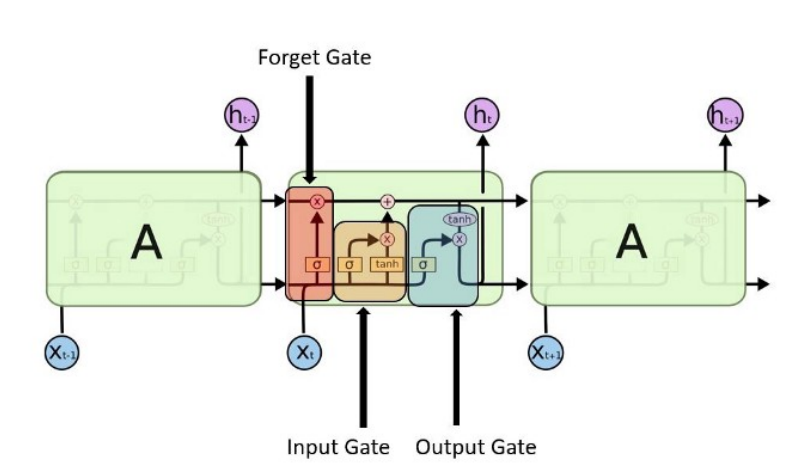
## Natural Language Processing

NLP atau *Natural Language Processing* adalah disiplin ilmu komputer yang bertujuan untuk memahami konsep dan maksud dari bahasa manusia. Sementara manusia cukup mahir memahami sintaks linguistik dan tata bahasa serta hubungan spasial tersirat, komputer memiliki kesulitan besar pengolahan query bahasa alami (Allen, 1995). Contoh sistem ini dan penggunaan API adalah *Natural Language Toolkit (*NLTK) dan CoreNLP.

Tujuan NLP adalah merancang dan membangun aplikasi yang memfasilitasi interaksi manusia dengan mesin dan perangkat lain melalui penggunaan bahasa alami.(Aprian Putra dan & Budi, n.d.)

## LSTM

LSTM *atau Long Short Term Memory* merupakan bentuk spesial dari RNN yang dapat melakukan pembelajaran pada dependensi jangka panjang (long-term 17 dependencies). Model ini diperkenalkan oleh Hochreiter dan Schmidhuber pada tahun 1997

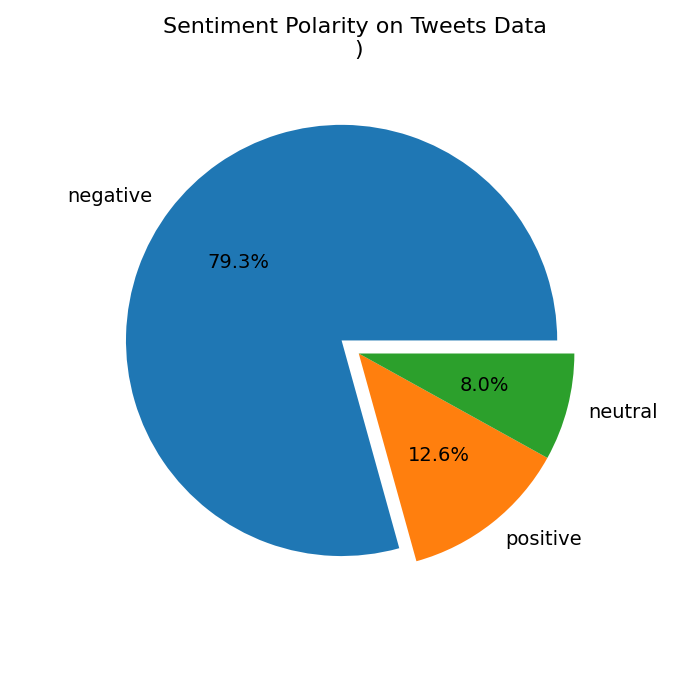


Gambar 2.1 Arsitektur LSTM

Pada *Long Short Term Memory* (LSTM) terdapat modifikasi pada RNN dengan menambahkan *memory cell* atau *memory unit* yang dapat menyimpan informasi yang dipelajari LSTM dalam jangka waktu yang panjang. LSTM memberikan solusi untuk mengatasi terjadinya *vanishing gradient* pada RNN saat memproses *data sequential* yang panjang.

## Sentimental Analysis

Sentimental Analysis adalah proses mendeteksi polaritas kontekstual teks. Yang berfungsi untuk menentukan apakah suatu teks tertentu positif, negatif, atau netral. Oleh karena itu sentimental analysis juga dapat dikatakan sebagai opinion mining, karena hasil yang diperoleh berasal dari opini atau kelakuan pembicara. Untuk analisis ini, opini yang dikumpulkan dari *user*, yang dapat digunakan untuk penelitian lebih lanjut. Media Sosial bertindak sebagai sebuah media dimana *user*nya dapat mengunggah banyak pendapat atau opini dalam sehari dan microblog tersebut dapat digunakan untuk klasifikasi. Banyak penelitian yang dilakukan di bidang sentimental analysis karena signifikansi dalam persaingan tingkat pemasaran dan perubahan kebutuhan masyarakat. Analisis sentimen memerlukan pemakaian seperangkat pelatihan untuk kinerjanya, dan kualitasnya memainkan peran penting dalam evaluasi teks yang akurat.



Gambar 2.2 Contoh diagram pada *Sentimental Analysis*

## Social Network Analysis

Social Network merupakan suatu jejaring yang menggambarkan struktur dalam situasi sosial jejaring sosial menggunakan nodes sebagai representasi dari manusia (atau sekelompok manusia) dan edges sebagai representasi dari interaksi sosial diantara dua manusia. Interaksi sosial ini dapat berupa hubungan pertemanan, hubungan kekeluargaan, hubungan professional, pertukaran barang, pertukaran uang, pola komunikasi, hubungan romantika, dan berbagai jenis hubungan lainnya.

Social Network Analysis (SNA) dapat dibentuk untuk merepresentasikan berbagai kondisi dan konteks dalam dunia nyata. Social Network direpresentasikan dengan menggunakan pendekatan teori graf, yang merupakan cabang dari ilmu matematika. Social Network dapat dibentuk ketika terdapat entitas yang berinteraksi dengan entitas - entitas lainnya dalam suatu lingkup sosial. Entitas ini dapat mewakili berbagai hal seperti manusia, perusahaan, ataupun organisasi.

SNA ini sering digunakan untuk menentukan aktor sentral di dalam sebuah jaringan dengan menghitung nilai centrality. (Tabassum et al., 2018). *Centrality* yang digunakan pada penelitian ini yaitu *degree centrality. Degree centrality* merupakan jumlah interaksi atau *edge* yang terjadi pada sebuah node tertentu. Rumus dari degree centrality adalah :

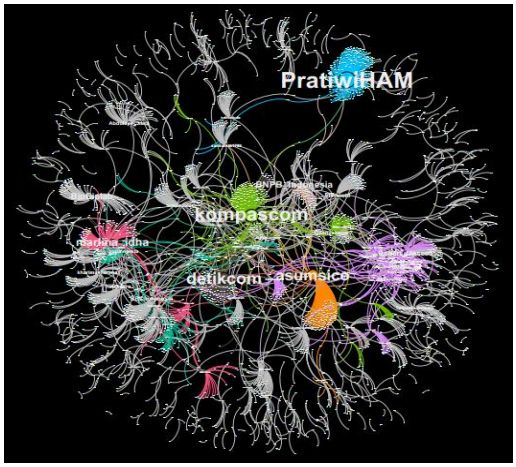
dimana:

CD = Degree centrality node i

aij = nilai baris ke-i dan kolom ke-j matriks adjacency

Seperti pada gambar 2.3, Label “PratiwiHam” termasuk dalam kluster yang cukup padat terlihat banyak node yang berkumpul di satu are membentuk kelompok. Setelah ditelusuri pada akun twitter yang bersangkutan ternyata merupakan akun dari pasien yang terkonfirmasi positif yang berhasil sembuh dari virus Covid-19 dan akun tersebut sering membuat thread berisi kisah hidupnya selama menderita Covid-19 yang mana threads yang dibuat oleh “@PratiwiHAM” mendapat banyak retweet dan likes dari pengguna twiter. (Inayah,dkk 2020)

Selain akun pengguna biasa, akun berita resmi seperti @detikcom, @kompascom dan @asumsico serta akun resmi pemerintah @BNPB\_Indonesia juga memiliki peran penting dalam penyebaran informasi terkait Covid-19 di twitter. Hal ini tidak terlepas dari seringnya akun-akun tersebut membuat tweet berkaitan dengan perkembangan dan informasi-informasi resmi dan terpercaya mengenai Covid-19. (Inayah,dkk 2020)



Gambar 2.3 Contoh implementasi Social Network Analysis pada penyebaran informasi covid-19 (Inayah,dkk 2020)

## Twitter

Twitter adalah layanan pengiriman pesan yang memiliki banyak kesamaan dengan alat komunikasi yang sekarang ini digunakan. Twitter memiliki elemen yang mirip dengan email, IM, SMS, blogging, RSS, dan sebagainya. Tetapi beberapa faktor, terutama kombinasi, membuat Twitter unik. Pesan yang dikirimkan dan diterima di Twittertidak lebih dari 140 karakter, yang artinya Twitter mudah digunakan untuk menulis dan membaca. Pesan-pesan di Twitter bersifat publik, yang artinya bertemu dengan orang baru di Twitter sangatlah mungkin terjadi. Walaupun sekarang ini ada fitur private yaitu dimana orang-orang yang ingin mengikuti kita harus mendapatkan izin dari kita untuk mereka bisa melihat kontenkonten Twitter milik kita, tapi berinteraksi dengan orang baru dengan platform ini sangatlah mungkin. Serta juga pesan yang dikirimkan dan diterima di Twitter sekarang ditambah menjadi 280 karakter. Pada situsnya, Twitter didefinisikan sebagai apa yang terjadi di dunia dan apa yang sekarang sedang orang-orang bicarakan, apa yang terjadi itu juga yang terjadi di Twitter, melihat apa yang orang-orang sedang bicarakan.

Twitter berulang kali mengubah kebijakan dalam persoalan API pada platformnya. Untuk saat ini per tanggal 9 Februari 2023 twitter mengeluarkan kebijakan baru yakni menghapus akses gratis dalam pengambilan data tweet baik itu API v1.1 dan v2.

## Text Mining

Text mining merupakan suatu teknik dalam ilmu komputer yang digunakan untuk memecahkan permasalahan informasi yang sangat banyak dengan mengkombinasikan teknik dari data mining Seperti pada data mining, text mining berusaha mengekstrak informasi yang bermanfaat dari sumber data melalui identifikasi dan eksplorasi pola menarik. Pada text mining sumber data adalah kumpulan dokumen, ini berarti dapat berupa koran, majalah, artikel, surat, ataupun laporan-laporan penelitian seperti jurnal , tugas akhir, atau tesis (Feldman dan Sanger, 2007). Dengan text mining dapat diketahui pengelompokkan dokumen yang serupa, memperkirakan penulis dari sebuah dokumen, atau mengklasfikasikan dokumen ke dalam suatu kategori. Tetapi tidak seperti data yang pada data mining yang umumnya sudah terstruktur, dokumen merupakan data yang tidak terstruktur. Proses text mining sama dengan tahapan Knowledge Discovery in Database (KDD) pada data mining(Widaningsih, 2018).

Berikut ini adalah proses text mining mengacu pada KDD antara lain:

a. Text Data

Dokumen adalah syarat utama dalam text mining. Data berupa teks merupakan suatu fragmen yang dianggap sebagai unit. Dapat berupa buku, paragraph, abstrak, atau pun judul. Untuk web, fragmen teks adalah halaman web.

b. Text PreProcessing

Pada tahap ini dilakukan pembersihan data dan tokenisasi. Pembersihan data dilakukan untuk menghilangkan informasi yang tidak perlu atau tidak diinginkan seperti tabel, gambar atau rumus-rumus. Pada tahap tokenisasi, dokumen diperlakukan sebagai bentuk string dan dipecah menjadi token. Pada token kalimat-kalimat yang ada pada dokumen dipisahkan setiap kata dan karakter yang membentuknya. Dengan membentuk token, maka dapat dilakukan analisis teks lebih lanjut (Bhumika, et al , 2013). Kata-kata yang bersifat stopword juga dihilangkan. Kata-kata stopword seperti “a”, “the” ,” but” adalah yang dibutuhkan pada struktur grammer dalam bahasa Inggris tetapi tidak memiliki arti dan tidak dibutuhkan dalam text mining . Pada tahap ini juga dilakukan proses stemming yang mengubah kata-kata dalam bentuk dasar. Misal : connection menjadi connect.

c. Feature Transformation (Attribute Generation)

Sebuah dokumen teks diwakili oleh kata-kata (fitur) yang dikandungnya dan kejadiannya. Representasi dokumen merupakan salah satu teknik yang digunakan untuk mengurangi kompleksitas dokumen sehingga lebih mudah untuk ditangani. Dokumen harus ditransformasikan dari versi teks lengkap ke dalam bentuk vector dokumen. Representasi dokumen yang umum digunakan disebut vector space model dimana dokumen direpresentasikan menjadi vektor dari kata-kata. Beberapa keterbatasannya adalah: representasi dimensi yang tinggi, hilangnya korelasi dengan kata-kata yang berdekatan dan hilangnya hubungan semantik yang ada didalam dokumen. Untuk mengatasi masalah ini, metode pembobotan kata digunakan untuk menetapkan bobot yang sesuai dengan kata tersebut. Pada model ruang vektor, dokumen diwakili oleh vektor kata-kata karena kata-kata yang membentuk dokumen akan menentukan isi dari dokumen tersebut (Bhumika et al, 2013).

d. Attribute Selection / Feature Selection

Tahap ini adalah teknik pengurangan dimensi yang efektif untuk menghilangkan fitur noise. Fitur noise merupakan informasi-informasi yang tidak berguna yang dapat mengganggu hasil penelitian dalam text mining, seperti tulisan copyright, menu navigasi pada halaman web, dan lain-lain (Ting et al, 2011).

Pemilihan fitur juga dikenal sebagai pemilihan variabel, adalah proses pemilihan subset dari fitur yang penting untuk digunakan dalam pembuatan model. Hal ini dilakukan karena data mengandung banyak fitur yang berlebihan atau tidak relevan. Misalkan fitur yang berulang adalah salah satu yang tidak memberikan informasi tambahan. Fitur yang tidak relevan tidak memberikan informasi yang berguna atau relevan dalam konteks apapun.. Pemilihan fitur dilakukan dengan menyimpan kata-kata dengan bobot tertinggi sesuai dengan ukuran yang telah ditentukan terhadap pentingnya kata tersebut.

e. Mining Pattern

Pada tahap ini teknik yang digunakan text mining digabungkan dengan teknik klasik pada data mining, seperti klasifikasi, pengelompokkan atau asosiasi. Teknik-teknik tersebut dapat mengolah data yang telah terstruktur sebagai hasil pengolahan dari tahapan text mining yang telah dilakukan.

f. Evaluation

Pada tahap ini dilakukan evaluasi dari hasil mining pattern. Hasil evaluasi umumnya menggunakan suatu nilai performansi sesuai dengan teknik data mining yang digunakan.

## Python

Bahasa python dibuat pertama kalinya oleh Guido Van Rossum. Bahasa ini termasuk dalam bahasa pemrograman tingkat tinggi yang mudah dipelajari karena struktur bahasa yang digunakan rapi dan mudah untuk dimengerti Tingkat fleksibilitas dan pengembangan python sesuai untuk digunakan dalam memecahkan berbagai masalah kompleks dengan lebih efisien termasuk artificial intelligence, machine learning, dan berbagai aktivitas data science lainnya. (Andry, 2020)

Beberapa pengguna menjalankan bahasa python hanya dari teks editor biasa namun untuk meningkatkan kemudahan banyak pengguna bahasa Python menggunakan suatu integrated development environment. (Andry, 2020)

Python diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif. Python dapat digunakan untuk berbagai keperluan pengembangan perangkat lunak dan dapat berjalan di berbagai platform sistem operasi.

Python didistribusikan dengan beberapa lisensi yang berbeda dari beberapa versi. Namun pada prinsipnya, python dapat diperoleh dan dipergunakan secara bebas, bahkan untuk kepentingan komersial. Adapun beberapa fitur yang diimiliki oleh python diantaranya sebagai berikut:

1. Memiliki kepustakaan yang luas. Dalam distribusi python telah disediakan modul-modul.
2. Memiliki tata bahasa yang mudah dipelajari.
3. Memiliki aturan layout kode sumber yang memudahkan pengecekan, pembacaan kembali dan penulisan ulang kode sumber. Berorientasi objek.
4. Dapat dibangun dengan bahasa python maupun C/C++. (Syahruddin, 2018)

## NLTK

NLTK (Natural Language Toolkit) adalah salah satu library Python yang paling banyak digunakan dalam Natural Language Processing. NLTK adalah library Python yang dapat memproses teks Natural Language dengan cepat dan mudah. Toolkit ini dikembangkan di University of Pennsylvania sebagai alat penelitian dan pengajaran untuk pemrosesan bahasa alami. NLTK memiliki sejumlah besar corpora built-in, termasuk berbagai jenis materi teks seperti novel, berita, teks obrolan jaringan, ulasan film, dll. termasuk Korpus Brown, Korpus Gutenberg, Korpus Alamat Pelantikan, Korpus Reuters, dll. (Kambhampati, 2019).

Selain itu, NLTK menyediakan antarmuka yang mudah digunakan ke lebih dari 50 kumpulan dan sumber daya leksikal seperti WordNet, bersama dengan rangkaian pustaka pemrosesan teks untuk klasifikasi, tokenisasi, stemming, tagging, parsing, dan penalaran semantik, pembungkus untuk perpustakaan NLP berkekuatan industri, dan forum diskusi aktif. Bekerja sama dengan Python yang kuat perpustakaan standar dan perpustakaan pihak ketiga lainnya, dapat melakukan pemrosesan sekunder dari hasil pemrosesan. Ini menyediakan dukungan yang kuat untuk memproses teks yang kompleks (Li, 2019)

## TwitterWebScraper

TwitterWebScraper adalah sebuah tools scraping yang di buat oleh WilfredNJH di github. Tools tersebut dibuat untuk mengumpulkan data twitter tanpa memerlukan API dari twitter. Namun tools tersebut memiliki kekurangan yang mana merujuk pada peraturan twitter mengenai Batasan melihat tweet per harinya. Output dari tools tersebut akan disimpan dengan format *file* *.csv*. tools tersebut menggunakan library selenium dan beautifulsoup4 untuk fungsi scraping tweetnya, dan untuk GUI nya menggunakan tkinter dan customtkinter. (Ng, 2023)

## NetworkX

Networkx adalah salah satu package pada bahasa pemrograman Python yang berfungsi untuk mengeksplorasi dan menganalisis jaringan dan algoritma jaringan. Networkx menyediakan berbagai jenis graf untuk menggambarkan jaringan, termasuk simple graph dan directed graph. Beberapa contoh algoritma graf yang dapat diimplementasikan dengan Networkx adalah shortest paths, betweenness centrality, closeness centrality, dan lain sebagainya.(Mufidah dkk, 2020)

## Visual Studio Code

Visual Studio Code adalah software yang dirilis oleh Microsoft sejak beralih ke pendekatan open source untuk .NET dan teknologi terkait. Microsoft juga telah mengembangkan software Visual Studio untuk mengintegrasikannya dengan framework JavaScript, manajer paket, dan perangkat UI. Model ASP.NET 5 menyederhanakan pengembangan web modern menggunakan kerangka kerja seperti Bootstrap, AngularJS, Knockout, Gulp, dan banyak lagi.

## Flask

Flask merupakan web framework yang ditulis dalam bahasa Python. Flask memiliki dua external libraries yaitu WSGI toolkit dan Jinja2 template engine. Flask merupakan jenis microframework yang tidak memerlukan library tertentu dalam penggunaannya. Flask dapat menggunakan ekstensi untuk menambahkan fitur dan komponen yang sudah disediakan oleh pihak ketiga dan tidak terpasang secara standar pada Flask seperti Form Validation, Upload Handling, dan Database(Aslam, 2015)

## NodeJs

Node.js merupakan sebuah perangkat lunak yang di desain untuk pengembangan perangkat lunak. Node.js sendiri dieksekusi pada sisi server. Platform ini menggunakan bahasa pemrograman JavaScript. Selain itu, Node.js juga menggunakan teknik nonblocking untuk mempercepat proses (Mubariz dkk., 2020).

## MySQL

MySQL merupakan software database open source yang sering digunakan untuk mengolah basis data yang menggunakan bahasa SQL (Subagia, 2018). MySQL adalah sebuah software database. MySQL merupakan tipe data relasional yang artinya MySQL menyimpan datanya dalam bentuk table-tabel yang saling berhubungan. Keuntungan menyimpan data di database adalah kemudahannya dalam penyimpanan dan menampilkan data karena dalam bentuk tabel. (Winarno, 2014). MySQL adalah salah satu aplikasi DBMS yang sudah sangat banyak digunakan para pemrogram aplikasi web. Kelebihan dari MySQL adalah gratis, handal, selalu di-update dan banyak forum yang memfasilitasi para pengguna jika memiliki kendala. MySQL juga menjadi DBMS yang sering di bundling dengan web server sehingga proses instalasinya jadi lebih mudah Priyanto, (Hidayatullah dkk, 2015).

## Metode Pengembangan Waterfall

System Development Life Cycle (SDLC) merupakan metodologi umum yang digunakan untuk mengembangkan sistem informasi. SDLC terdiri dari beberapa fase yang dimulai dari fase perencanaan, analisis, perancangan, implementasi hingga pemeliharaan sistem. Metode waterfall adalah salah satu model SDLC yang sering digunakan dalam pengembangan sistem informasi atau perangkat lunak. Model waterfall menggunakan pendekatan sistematis dan berurutan. Tahapan model waterfall antara lain requirement, design, implementation, verification,dan maintenance. Kelebihan menggunakan metode waterfall dalam pengembangan sistem informasi adalah kualitas dari sistem yang dihasilkan akan baik karena pelaksanaannya dilakukan secara bertahap, sementara untuk kekurangannya adalah proses pengembangan sistem membutuhkan waktu yang lama sehingga biaya yang diperlukan juga mahal. Metode waterfall cocok digunakan untuk proyek pembuatan sistem baru dan juga pengembangan sistem atau perangkat lunak yang berskala besar.(Aceng Abdul Wahid, 2020)

## Pengujian Black Box

Black box testing merupakan pengujian kualitas perangkat lunak yang berfokus pada fungsionalitas perangkat lunak. Pengujian black box bertujuan untuk menemukan fungsi yang tidak benar, kesalahan antarmuka, kesalahan pada struktur data, kesalahan performansi, kesalahan inisialisasi dan terminasi (Ningrum dkk., 2019)

Pengujian Black Box bertumpu pada memastikan tiap proses sudah berfungsi sesuai dengan kebutuhan yang diharapkan. Penguji dapat mengartikan himpunan kondisi masukan dan menjalankan pengujian pada pengkhususan fungsi dari sistem. Sehingga pengujian merupakan suatu cara pelaksanaan program yang bertujuan menemukan kesalahan atau error kemudian memperbaikinya sehingga sistem dapat dikatakan layak untuk digunakan. Blackbox testing terbagi atas lima tahapan, yaitu: Graph based testing, Equivalence partitioning, Boundary value analysis, Comparison testing, Orthogonal array testing.

# METODE PENELITIAN

## Tempat dan Waktu Penelitian

Penelitian ini akan mulai dilaksanakan bertempat di Politeknik Negeri Ujung Pandang dengan estimasi waktu pengerjaan selama 7 bulan mulai dari bulan Januari s/d Julii 2023.

## Alat dan Bahan

Adapun alat dan bahan yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Hardware

Komputer dengan spesifikasi Intel Core i3 dengan ram 8 GB.

1. Software
2. Sistem Operasi Windows 10
3. Visual Studio Code
4. NodeJS
5. Library Pandas, Scikitlearn, Tensorflow, NLTK, Twint, NetworkX, Flask
6. MySQL

## Prosedur Penelitian

Dalam pembuatan penelitian ini, terdapat beberapa tahapan agar metodologi dan perancangan sistem tercapai. Tahap-tahap yang dilaksanakan adalah sebagai berikut:

### **Metode Penelitian dan Tahapan Pengembangan Aplikasi**

Perancangan sistem menggunakan bagan alur atau flowchart. Flowchart akan dibagi menjadi beberapa, yaitu flowchart yang menggambarkan cara kerja aplikasi secara umum dan beberapa flowchart yang akan menjelaskan lebih detail.

Chart, diagram

Description automatically generated

Gambar 3.1 Metode waterfall

### **Analisis Kebutuhan**

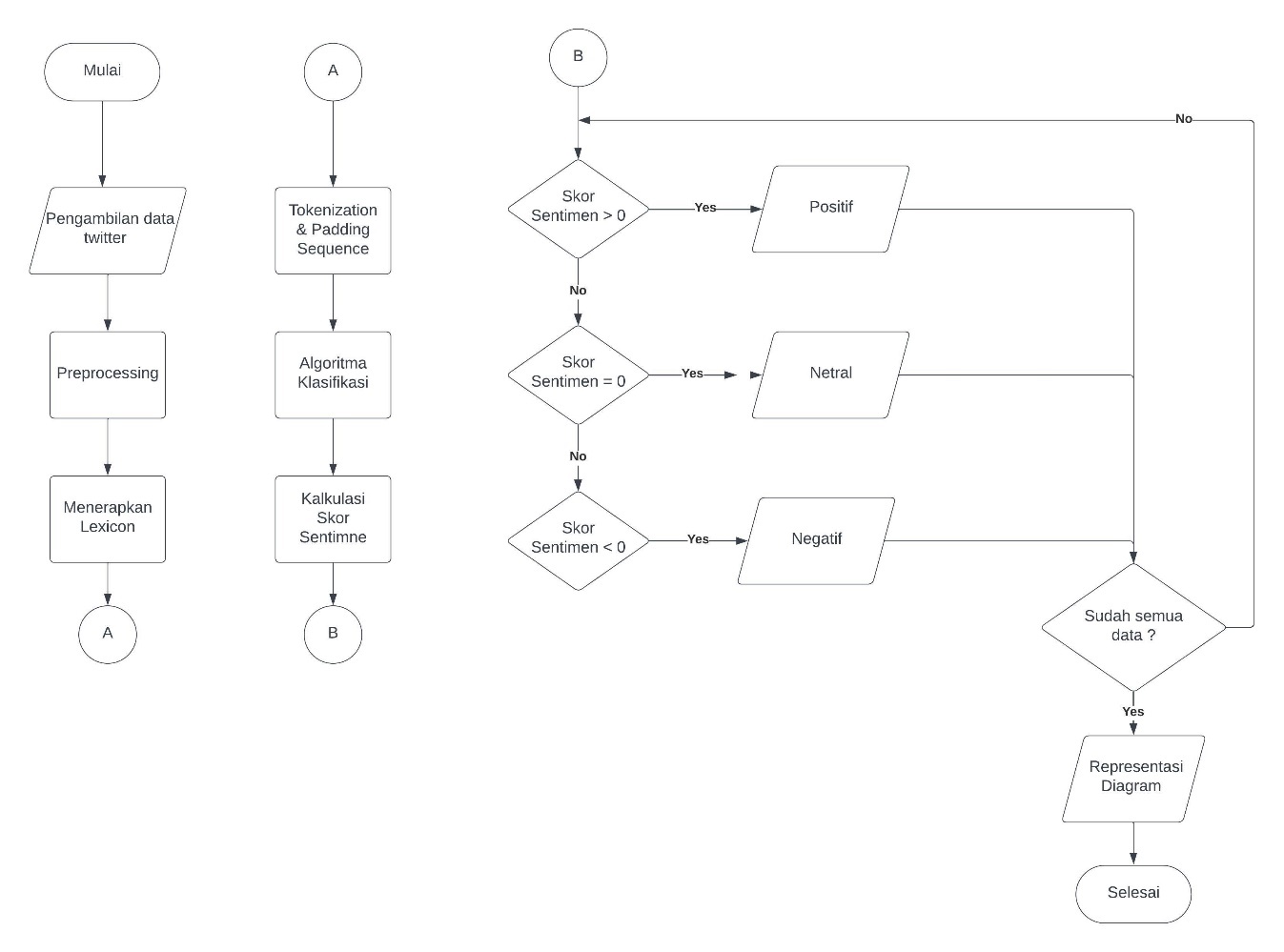
Tahap ini adalah tahap awal dimana dilakukan analisis terhadap kebutuhan apa saja yang diperlukan saat proses pengerjaan berlangsung. Adapun untuk kebutuhan pada proses pengambilan data melibatkan tools TwitterWebScraper yang dimana akan menyimpan file data ke dalam format csv yang berisikan kolom (username, handle, postDate, responding, retweet, dan likes).

### **Desain Sistem**

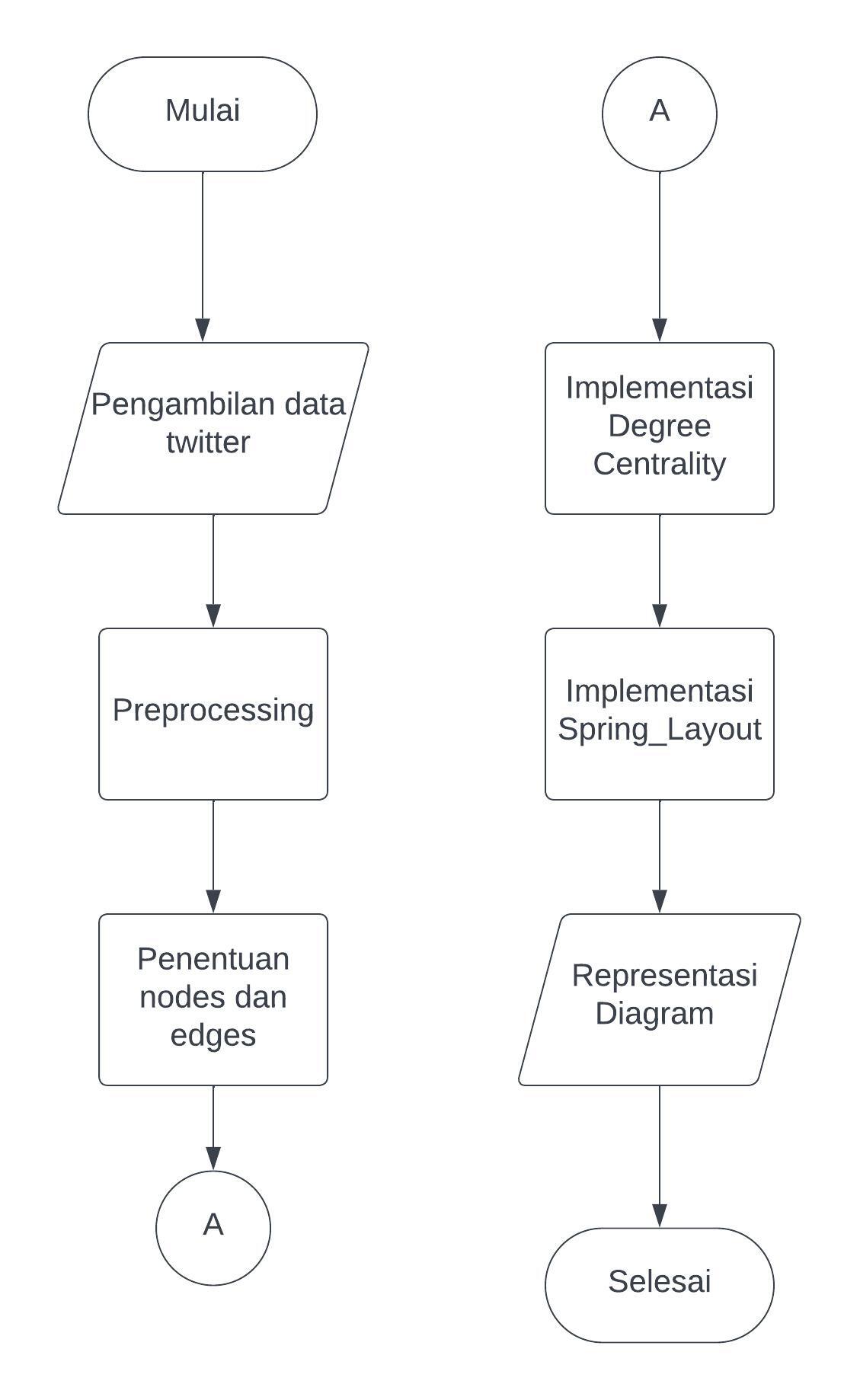
Dalam tahap desain pada penelitian ini menggunakan beberapa pemodelan yaitu: flowchart, use case diagram, dan activity diagram. Untuk desain user interface pada penelitian ini penulis menggunakan pemodelan wireframe.

#### Flowchart

Flowchart berfungsi untuk mengetahui setiap proses yang terjadi. Dalam penelitian ini, penulis membuat 2 flowchart dikarenakan ada 2 proses yang berbeda yaitu flowchart untuk proses sentimental analysis dan flowchart untuk proses network analysis.



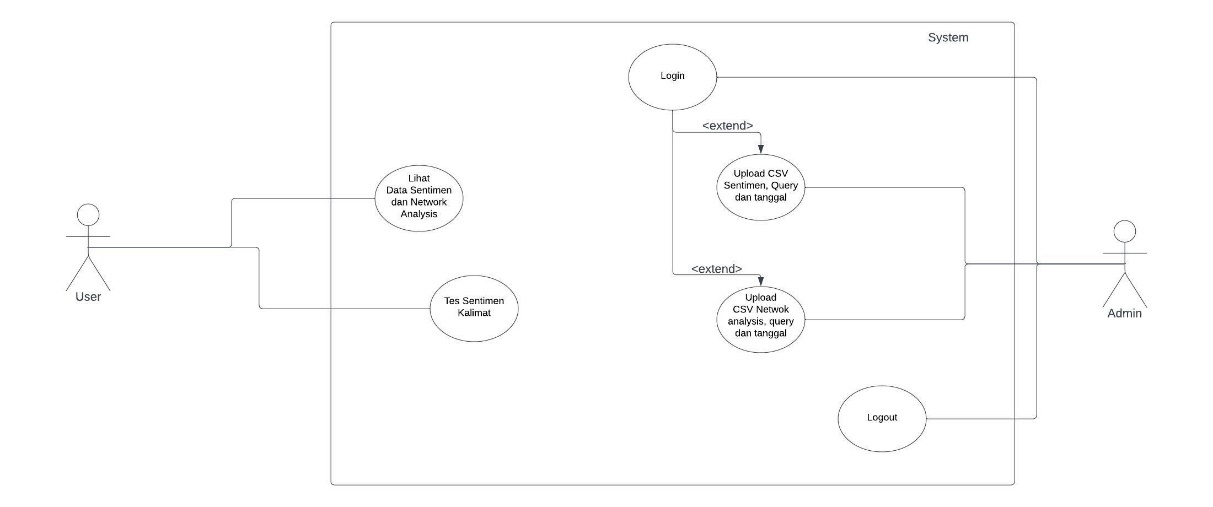
**Gambar 3.2 Tampilan flowchart sentiment analysis**



**Gambar 3.3 Flowchart Network Analysis**

#### Use Case Diagram

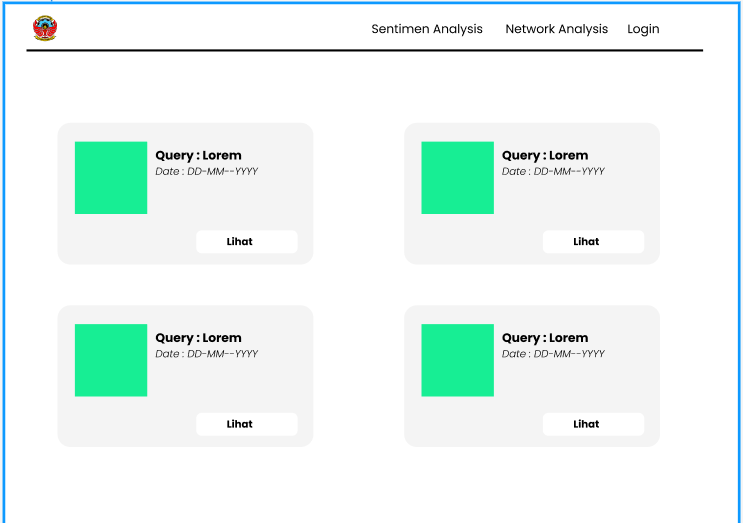
Use case diagram disini berfungsi untuk memperlihatkan fase awal tiap proses dalam sistem yang diterapkan pada penelitian ini. Dimana hal tersebut dapat memudahkan dalam tahap pengembangan nantinya.



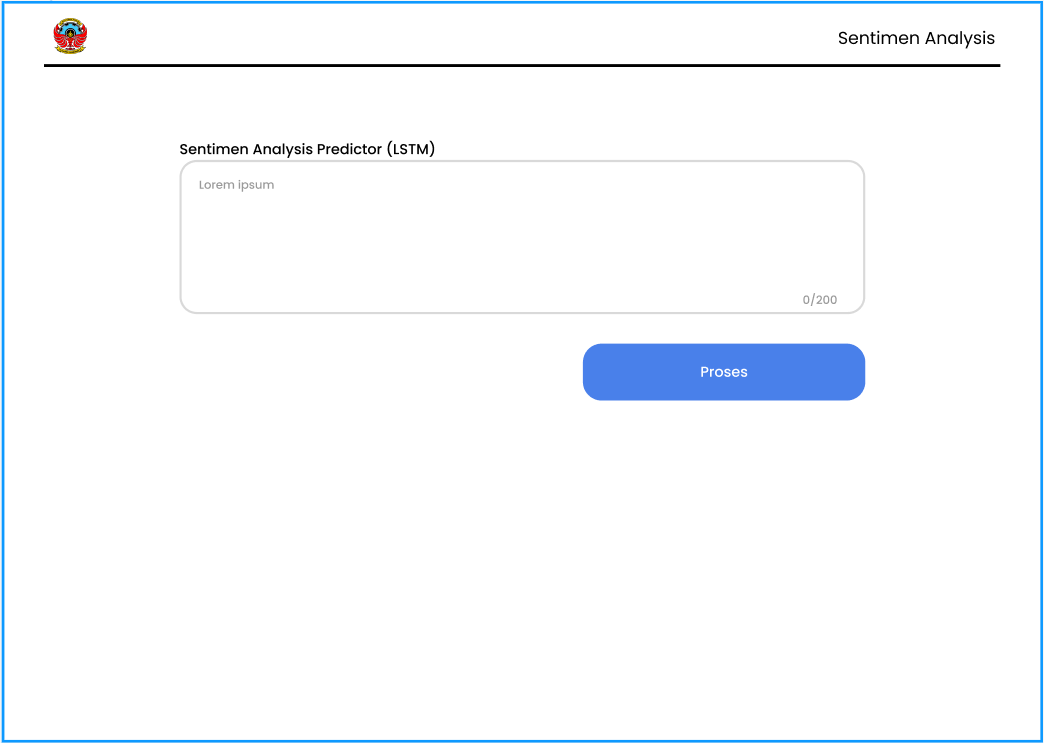
**Gambar 3.4 Use Case Diagram**

#### Wireframe

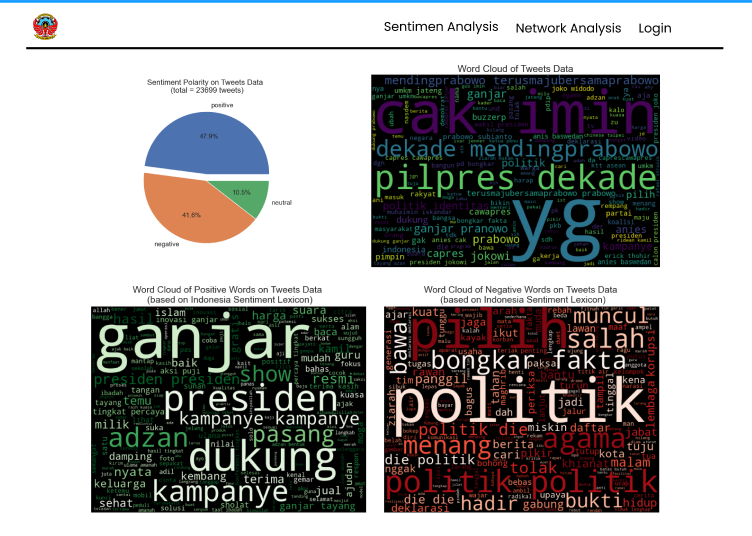
Wireframe adalah sebuah desain kerangka tahap awal yang menggambarkan struktur yang akan dibuat nantinya. Desain wireframe tersebut nantinya akan dilanjutkan ke tahap pembuatan desain user interface nantinya.

****

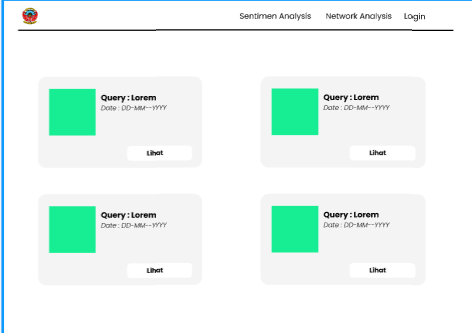
**Gambar 3.5 Tampilan halaman home**

****

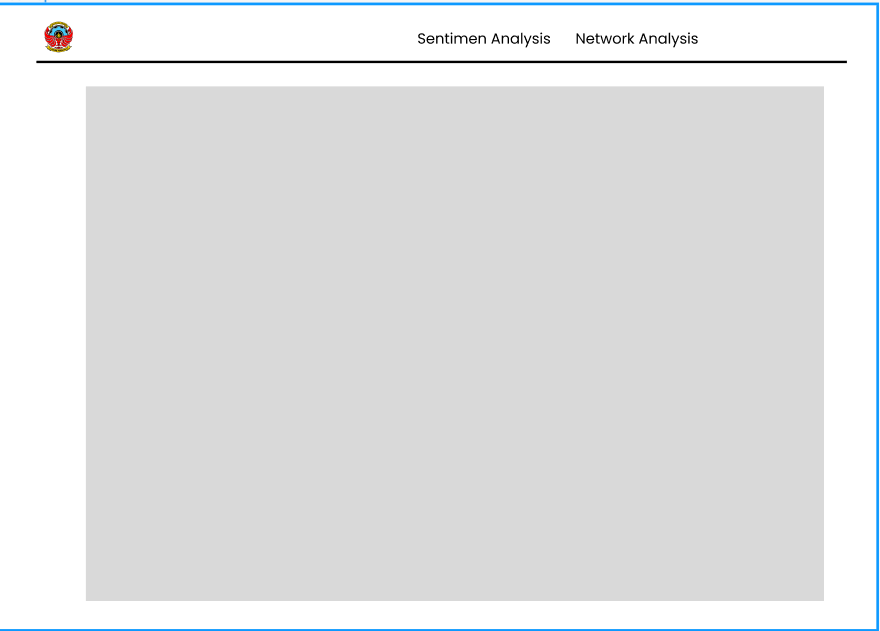
**Gambar** **3.6 Tampilan halaman Analisa sentimen**

****

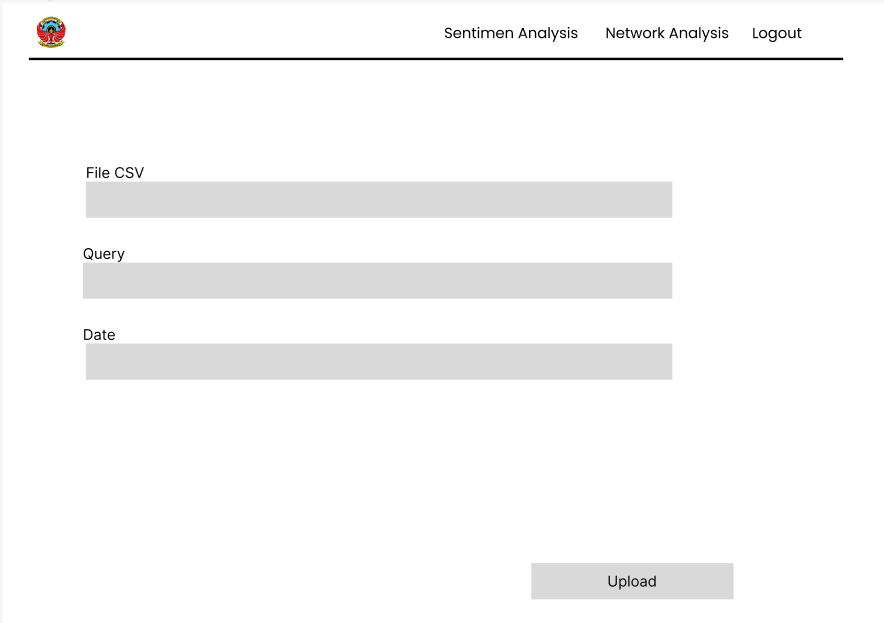
**Gambar 3.7 Tampilan halaman visualisasi jika card lihat ditekan**

****

**Gambar 3.8 Tampilan halaman Network Analysis**

****

**Gambar 3.9 Tampilan halaman jika card pada Network Analysis ditekan**

****

**Gambar 3.10 Tampilan halaman admin**

### **Implementasi dan Pengkodean**

Dalam tahap ini penelitian mulai dibangun sesuai dengan analisis kebutuhan. Untuk pengkodeannya akan menggunakan Bahasa python dengan IDE google colaboratory dan Visual Studio Code

### **Testing**

Pada tahapan ini pengujian program dilakukan dengan menggunakan Blackbox Testing dengan harapan bahwa perancangan yang sudah dibuat dapat berjalan dengan sesuai kehendak.

### **Pemeliharaan**

Dalam proses pemeliharaan ini penulis mengupayakan pengembangan sistem yang telah di rancang terkait software untuk dapat dibuat maksimal agar aplikasi dapat berjalan dengan baik.

## Pengujian Sistem

Pengujian merupakan tahapan yang cukup berperan penting didalam sebuah perancangan dan pengembangan aplikasi. Adapun pengujian yang akan digunakan adalah menggunakan metode Black-Box Testing. Black-Box Testing adalah pengujian yang dilakukan untuk melihat hasil keluaran atau output dari aplikasi yang telah dibangun. Pada metode Black-Box ini akan dilakukan uji data dan pemeriksaan terhadap fungsional dari aplikasi (Adytio, 2017). Metode ini memfokuskan pada fungsional dari suatu sistem aplikasi.

Pengujian dalam *blackbox testing* ini lebih mengaju kepada bagaimana user atau dapat dikatakan pengguna akhir dapat melakukan interaksi pada system aplikasi yang dibuat.