SENTIMEN ANALISIS PILPRES 2024 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN NAÏVE BAYES CLASSIFIER

PROPOSAL SKRIPSI

Karya tulis sebagai salah satu syarat
untuk memperoleh gelar Tingkat Sarjana

Oleh

MUCHAMMAD FAHD ISHAMUDDIN
NPM : 411550050180048

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS LANGLANGBUANA
2023

LEMBAR PENGESAHAN PROPOSAL

<table>
<tr><td>JUDUL</td><td colspan="2">SENTIMEN ANALIS PILPRES 2024 PADA MEDIA SOSIAL TWITTER MENGGUNAKAN NAÏVE BAYES CLASSIFIER</td></tr>
<tr><td>PENGUSUL
Nama:
NPM :</td><td colspan="2">MUCHAMMAD FAHD ISHAMUDDIN
411550050180048</td></tr>
<tr><td rowspan="3"></td><td colspan="2">DISETUJUI sebagai SKRIPSI di Program Studi Teknik Informatika oleh:</td></tr>
<tr><td></td><td></td></tr>
<tr><td>Reviewer
Nama:
NIDN:</td><td>Arief Ginanjar,S.T., M.Kom.
0423107805</td></tr>
<tr><td></td><td colspan="2">Bandung, _____

Ketua Program Studi
Teknik Informatika

<u>(Yiti Supendi, S.Kom., M.T.)</u>
NIDN: 0415046303</td></tr>
</table>

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

BAB I PENDAHULUAN

I.1 Latar Belakang

Indonesia merupakan negara yang memiliki bentuk pemerintahan presidensial dan demokrasi. Pemerintahan presidensial berarti kepemimpinan pada negara tersebut dipimpin oleh seorang presiden, Demokrasi berarti kekuasaan tertinggi ada di tangan rakyat sehingga yang dapat memilih siapa pemimpin pada negara tersebut. Seperti yang kita ketahui jika pemilihan presiden diadakan dengan Pemilu(Pemilihan Umum) yang bertujuan untuk menentukan eksekutif dan legislatif serta diselenggerakan oleh KPU(Komisi Pemilihan Umum).

Pemilu pertama kali dilaksanakan pada 29 September 1955 untuk memilih anggota DPR dilanjutkan pada 15 Desember 1955 untuk memilih anggota Dewan Konstituante, pada saat ini Indonesia masih dipimpin oleh Ir. Soekarno. Pada tahun 1967 ada perjanjian SUPERSEMAR yang menyatakan penyerahan kepemimpinan dari Ir. Soekarno kepada Soeharto. Pemilu pada tahun 1971,1977,1997 ketiga pemilu tersebut hanya digunakan untuk memilih DPR hingga akhirnya pada tahun 1999, setelah masa kepemimpinan Ir. B. J. Habibie, FREng. Diadakan pilpres tetapi melali siding paripurna MPR,dengan mencatatkan Abdurrahman Wahid (gus dur) menjadi presiden ke-4 Indonesia dan diampingi Megawati Soekarnoputri sebagai wakil presiden, hingga pada tahun 2004 merupakan pertama kalinya pemilihan presiden dilakukan secara luberjurdil (langsung, umum, bebas, rahasia, jujur dan adil) dengan memenangkan Susilo Bambang Yudhoyono sebagai presiden dan Jusuf Kalla sebagai wakil presiden. Terakhir dilaksanakn pada 2019 yang dimenangkan oleh petahan yakni Ir. H. Joko Widodo.

Hingga sekarang pemilihan presiden masih dilakukan untuk menentukan pemimpin negara Indonesia, tetapi ada yang berbeda antara zaman terdahulu dengan zaman sekarang, pada zaman sekarang kita bisa mengetahui reaksi masyarakat terhadap pemilu dan juga calon presiden

yang diusung, apalagi melalui social media yang merupakan komponen primer manusia era modern.

Gambar 1. 1 Grafik Pengguna Aktif Media Sosial di Indonesia

Pengguna media social di Indonesia per tahun 2022 dapat dilihat pada gambar 1.1 di atas dan menunjukkan sudah mencapai 191 juta jiwa dan total masyarakat Indoneisa ialah 275 juta jiwa, berarti sudah mencapai 69% jiwa di Indonesia menggunakan media social hal ini bisa menjadi alasan media social adalah media yang bisa menjadi pertimbangan dalam melihat elektabilitas maupun perspektif masyarakat terhadap Pemilu tersebut. Media social yang sering menjadi persebaran opini dari masyarakat Indonesia ialah twitter, karena banyaknya pendapat yang disampaikan dari warganet Indonesia akan menghasilkan berbagai macam reaksi, maka dapat dilakukan sentiment analisis untuk mendapatkan nada emosional tweet warganet Indonesia, algoritma yang banyak digunakan untuk mendapatkan sentiment ialah naïve bayes, dengan menghitung probabilitas dengan dasar bayes theorem.

Setelah mendalami permasalahan tersebut, maka penulis tertarik untuk menganalisa data dari twitter tentang Pilpres 2024 dan menuangkannya pada penelitian Skripsi yang berjudul ”**SENTIMEN ANALISIS PILPRES 2024 PADA MEDIA SOSIAL**

TWITTER MENGGUNAKAN NAÏVE BAYES CLASSIFIER"

I.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah diketahui, maka yang menjadi rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana dapat menilai pendapat yang disampaikan oleh warganet pada media social Twitter?
2. Bagaimana membuat tweet yang dapat menaikkan sentiment pemilu?

I.3 Batasan Masalah

Adapun Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data yang digunakan ialah data scrapping yang berasal dari media social twitter sebanyak 93882 data
2. Jangka waktu data yang digunakan ialah data tweet yang terjadi pada bulan januari hingga Desember 2022
3. Output hanyalah data sentiment terhadap Pilpres 2024

I.4 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka yang menjadi tujuan penelitian yang akan dilakukan antara lain:

1. Menerapkan metode naïve bayes classifier guna melakukan sentiment analisis dari social media twitter agar menjadi pertimbangan kontestan politik.
2. Mengetahui akurasi terbaik dan melakukan sentiment analysis tentang pilpres 2024 guna mengetahui bentuk opini positif, negative dan netral.

I.5 Keluaran Penelitian

Dalam penelitian ini luaran yang dihasilkan adalah sebagai berikut:

1. Python Notebook dan visualisasi data

2. Laporan Penelitian

I.6 Rencana Kegiatan

Tabel 1.1 Rencana Kegiatan

No	Uraian Kegiatan	Bulan 1				Bulan 2				Bulan 3			
1	Kebutuhan Sistem												
2	Penulisan Bab I												
3	Pendahuluan												
4	Penulisan Bab II												
5	Landasan Teori												
6	Penulisan Bab III												
7	Metodologi Penelitian												
8	Penulisan Bab IV												
9	Hasil dan Pembahasan												
10	Penulisan Bab V												
11	Kesimpulan dan Saran												
	Tahapan Utama Kegiatan												

BAB II LANDASAN TEORI

II.1 Teori Terkait Permasalahan

II.1.1 *Sentiment Analysis*

Analisis sentimen adalah pengumpulan pandangan orang tentang setiap peristiwa yang terjadi dalam kehidupan nyata. Dalam situasi seperti itu di mana dunia sedang melalui, memahami emosi dari orang-orang berdiri sangat penting. Skenario kubur dimana orang tidak bisa keluar dari rumah mereka menuntut eksplorasi-ing apa orang-orang benar-benar berpikir tentang keseluruhan skenario. Oleh karena itu, penulis telah merencanakan pekerjaan ini di bawah menghadapi situasi yang menuntut terutama di media social (Chakraborty,K. 2020).

Analisis sentimen adalah proses untuk mengidentifikasi dan mengenali atau mengkategorikan emosi pengguna atau pendapat untuk layanan apa pun seperti film, masalah produk, acara, atau setiap atribut adalah positif, negatif atau netral. Sumber untuk analisis ini adalah saluran komunikasi sosial yaitu situs Web yang meliputi *review*, forum diskusi, *blog*, *micro-blog*, *Twitter* dll. Bidang penelitian ini sangat populer saat ini karena data pendapatnya di mana pengguna dapat menemukan ulasannya layanan apa pun yang berguna untuk kehidupan sehari-hari mereka. Besar jumlah data opini disimpan dalam bentuk digital. Untuk topik tertentu atau pendapat analisis sentimen yang menghubungkan penambangan data bekerja dan memberikan output. (Mehta, P. and Pandya, S., 2020)

II.1.2 Naïve Bayes Classifier

Naïve Bayes Classifier adalah metode klasifikasi berdasarkan teorema Bayes. Pengklasifikasi *Naïve Bayes* dikenal lebih baik daripada beberapa metode klasifikasi lainnya. Karena pertama, ciri utama dari *Naïve Bayes* adalah asumsi independensi (naif) yang sangat kuat dari setiap kondisi atau peristiwa. Kedua, modelnya simple dan mudah dibuat. Ketiga, model dapat diimplementasikan untuk set data yang besar. Dasar salah satu teorema *Naïve Bayes* yang digunakan adalah rumus Bayes sebagai berikut: (Han, Kamber, & Pei, 2012).

Metode Naïve bayes classifier berasal dari bayes theorem yang ditemukan oleh Thomas bayes pada tahun 1770. Teorema bayes adalah sebuah teorema dengan

dua penafsiran berbeda. Teorema ini menyatakan seberapa jauh derajad kepercayaan subjektif harus berubah secara rasional ketika diberikan petunjuk baru. Teori ini juga berasal dari penerapan teori probabilitas.

Berikut rumus dari teori naïve bayes:

$$P(H|e) = \frac{P(e|H)P(H)}{P(e)}$$

P(H|e) = peluang kejadian **H** apabila **e** terjadi

P(e|H) = peluang kejadian **e** apabila **H** terjadi

P(H) = probabilitas kejadian (H)

P(e) = probabilitas (e) atau disebut prior probability. Berlaku jika (e) ≠ 0

II.1.3 Machine Learning

Machine learning atau dalam Bahasa Indonesia dikenal dengan pembelajaran mesin adalah aplikasi dari disiplin ilmu kecerdasan buatan (*Artificial Intelligence*). Konsep dari machine learning adalah memberikan kemampuan kepada computer untuk belajar secara mandiri dari sekumpulan data yang sudah diberikan sebelumnya, dengan menggunakan algoritma dan model untuk membuat prediksi. Fokus utama dari machine learning adalah untuk menemukan sebuah pola yang tepat dari sekumpulan data, sehingga dapat menghasilkan suatu model untuk melakukan proses *input-output* tanpa menggunakan kode program secara eksplisit (A. K. Tiwari, 2017). *Machine learning* dibagi dalam 3 bentuk, yakni *supervised learning*, *unsupervised learning* dan *generative learning*. Sentiment analisis dengan algoritma naïve bayes menggunakan metode *supervised Learning*.

1. *Supervised Learning*

Supervised learning adalah bidang pengenalan pola dan statistik dalam ilmu komputer. Ini adalah studi ilmiah tentang algoritma dan model statistik, yang digunakan untuk melakukan tugas tertentu secara efisien, tanpa menggunakan instruksi eksplisit, tetapi mengandalkan model. Algoritme pembelajaran yang diawasi membangun model matematika dari data sampel untuk membuat prediksi tanpa

memerlukan pemrograman eksplisit untuk melakukan tugas. (Yin, Q. 2020).

Supervised Learning adalah suatu metode untuk menciptakan *artificial intelligence* (AI), untuk mengidentifikasi pola dalam kumpulan data yang tidak di klasifikasikan atau tidak di beri label. Algoritma yang bertujuan untuk memperkirakan fungsi pemetaan sehingga ketika ada variabel *input* (X) kita dapat memprediksi variabel *output* (Y). Algoritma supervised learning dapat digunakan untuk memproses berbagai jenis data, mulai data yang terstruktur hingga yang tidak terstruktur. (Altamevia, F. 2023)

II.2 Teori Pendukung

II.2.1 Python

Python adalah bahasa pemrograman komputer open source untuk tujuan umum. Ini dioptimalkan untuk kualitas perangkat lunak, produktivitas pengembang, portabilitas program, dan integrasi komponen. Python digunakan oleh setidaknya ratusan ribu pengembang dunia di berbagai bidang seperti skrip Internet, pemrograman sistem, antarmuka pengguna, kustomisasi produk, pemrograman numerik, dan banyak lagi. Secara umum dianggap menjadi salah satu dari empat atau lima bahasa pemrograman yang paling banyak digunakan di dunia hari ini.(Mark Lutz, 2011)

Python adalah bahasa pemrograman interpretative yang dianggap mudah dipelajari serta berfokus pada keterbacaan kode, dengan kata lain python diklaim sebagai Bahasa pemrograman yang memiliki kode kode pemrograman yang sangat jelas, lengkap dan mudah untuk dipahami. (JUD, 2019)

II.2.2 Text Mining

Text mining adalah salah satu bidang yang sampai saat ini masih berkembang dengan pesat, dengan tugasnya dalam mengekstraksi atau mengumpulkan informasi yang bermakna dari teks alami suatu bahasa. Ini dapat diartikan sebagai proses menganalisis suatu teks untuk kemudian diekstrak informasi-informasi yang berguna dari teks tersebut untuk tujuan terentu. Dalam

budaya modern, teks adalah salah satu media dalam pertukaran informasi, dibanding dengan database, teks tidak terstuktur, memiliki bermacam-macam bentuk, dan lebih sulit ditangani menggunakan algoritma tertentu.(Witten : 2004)

Pada kasus ini, yaitu text mining sumber data yang berupa teks tidak memiliki struktur yang jelas dan memiliki bermacam bentuk, sehingga disebut sebagai unstructured data. Maka dari itu, butuh proses untuk membuat data menjadi lebih terstuktur sehingga ekstraksi informasi dari teks akan lebih mudah, tepat, dan sangat penting dalam proses text mining. Sumber data yang digunakan, yaitu novel berbahasa Indonesia merupakan unstructured data, sehingga butuh proses untuk membuat data menjadi lebih terstuktur. Salah satunya adalah dengan diawali oleh preprocessing, yang mana nanti akan menghasilkan fitur yang lebih represenatif dibanding sumber data novel berbahasa Indonesia yang belum dipersiapkan dan masih tidak berstruktur.

II.2.3 Jupyter Notebook

Jupyter adalah organisasi non-profit untuk mengembangkan software interaktif dalam berbagai bahasa pemrograman. Notebook adalah satu software buatan Jupyter, adalah aplikasi web open-source yang memungkinkan Anda membuat dan berbagi dokumen interaktif yang berisi kode liv e, persamaan, visualisasi, dan teks naratif yang kaya.(B. Prijono : 2019)

Pada penelitian ini, Jupyter Notebook digunakan sebagai salah satu text editor untuk menuliskan kode-kode program Python serta memvisualisasikan data hasil olah pada sistem ini.

II.2.4 Preprocessing

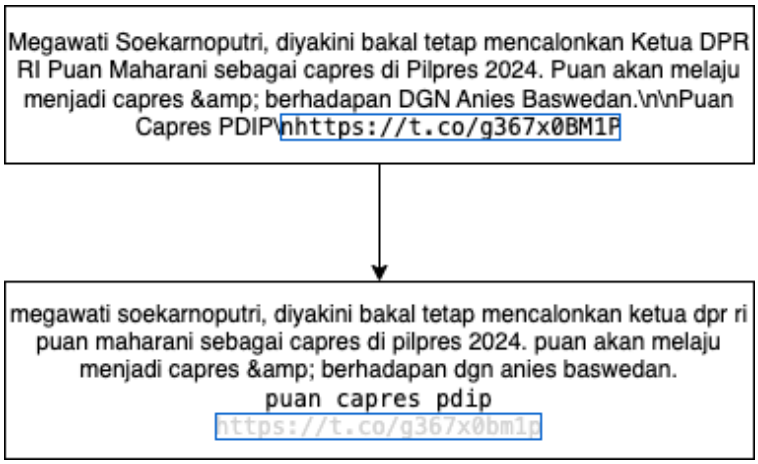
Proses preprocessing merupakan tahap dimana sumber data input diolah Kembali sebelum kemudian diproses lebih lanjut dan dijadikan bahan data pada machine leardning. Pada kasus teks tweet, text mining masih memliliki struktur yang bercampur dimana pada tweet masih ada mention dan link, serta data yang NaN pada data hasil scrapping, sehingga dibutuhkan proses yang merubah bentuknya menjadi data yang terstruktur. Proses ini akan melakukan penyeragaman case tweet yang merubah semua tweet menjadi lowercase, menghilangkan tanda mention serta

username termention, menghilangkan tautan pada tweet, kemudian membuat token dari data input, sehingga data lebih bersih, terstruktur dan dapat diolah lebih lanjut.

Pada penelitian ini, tahap preprocessing yang diterapkan adalah *Case Folding, Lemmatization, Stopword Removal dan Tokenizing.*

1. Case Folding

Case folding digunakan untuk menyeragamkan seluruh tesx dalam case yang seragam, abik menjadi huruf kecil (lowercase) ataupun huruf kapital (uppercase). Case folding digunakan pada penelitian ini adalah penyeragaman menjadi lowercase. Contoh dari penerapan case folding bisa dilihat pada gambar dibawah ini:

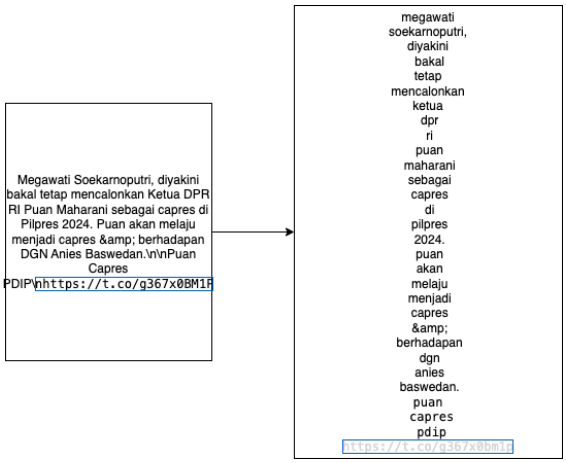


Gambar 2. 1 contoh case folding

2. Tokenizing

Proses Tokenizing adalah proses dimana string dipotong menjadi beberapa bagian dengan melihat delimiternya, seperti tipe kapitalisasi, keberadaan digit, tanda baca, karakter special dan sebagainya. Pemecahan dokumen menjadi kata – kata tunggal silakukan dengan caara men-scan dokumen dan setiap kata akan teridentifikasi atau terpisahkan dengan kalimatnya oleh delimiter. Tokenizing adalah proses diamana data input dibagi menjadi beberapa token sesuai dengan jumlah kalimat

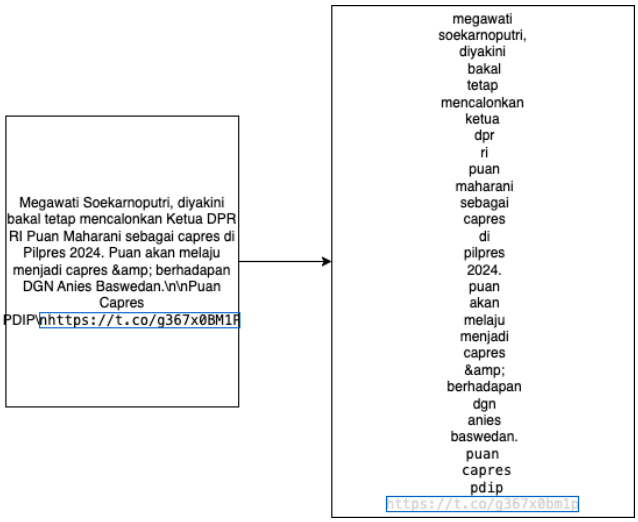
menggunakan delimiter '.' Pada teks input contoh tokenizing ada pada gambar dibawah.



Gambar 2. 2 contoh tokenizing

3. Stemming

Stemming adalah teknik pada natural language processing yang digunakan untuk mengembalikan kata kepada kata dasarnya yang disesuaikan dengan kamus Bahasa Indonesia, proses stemming dilakukan dengan menggunakan library sastrawi. Stemming digunakan pada kebutuhan yang berhubungan dengan text mining seperti information retrieval yang dilakukan pada tahap preprocessing.



Gambar 2. 3 contoh stemming

4. Stopword Removal

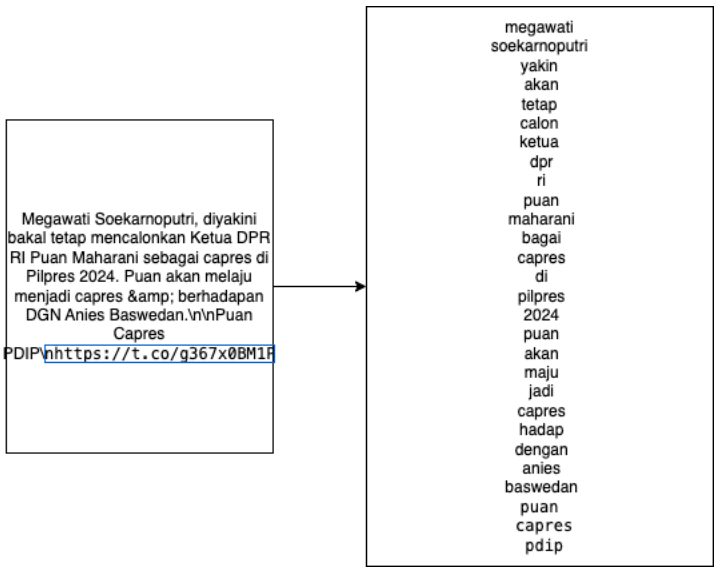
Stopword Removal adalah tahap pemilihan kata-kata yang dianggap penting. Terdapat dua metode yang dapat digunakan dalam tahap stopword removal, yakni:

- **Stoplist**

  Pada metode ini, kita menyuaokan kumpulan kata yang tidak deskriptif/tidak pendting yang disebut stoplist. Kata yang termasuk ke dalam stoplist akan dibuang dan tidak digunakan pada proses selanjutnya.

- **Wordlist**

  Wordlist merupakan kebalikan dari stoplist, pada metode ini kita menyiapkan kumpulan kata yang deskriptif yang disebut wordlist. Hanya kata yang termasuk ke dalam wordlist yang akan digunakan pada proses selanjutnya, sementara kata lainnya akan dibuang



Gambar 2. 4 contoh stopword removal

II.3 Penelitian-penelitian Terdahulu

II.3.1 Implementasi Metode Naïve Bayes untuk Analisis Sentimen Warga Jakarta Terhadap Kehadiran Mass Rapid Transit

Sarika A, Helena N, Noor F, Ika N.(2019), melakukan penelitian menggunakan data dari sosial media yaitu Twitter dengan keyword "MRTJakarta" yang dilakukan selama masa uji coba public MRT yaitu dari tanggal 5 – 23 maret 2019. Tweet yang diambil sebanyak 1000 tweet (800 tweet untuk training dan 200 tweet untuk testing). Dalam peneliltian ini naive bayes dapat memprediksi sentimen dari tweet yang sudah dikumpulkan terkait animo masyarakat terhadap MRTJakarta dengan akurasi sebesar 75%.

II.3.2 Algoritma Naïve Bayes Classifier Untuk Analisis Sentiment Pengguna Twitter Terhadap Provider By.u

Ike V, Bagas S.(2022), melakukan penelitian ini dapat diambil kesimpulan bahwa Algoritma Naïve Bayes Classifier dapat melakukan analisis sentimen dengan benar dan melakukan klasifikasi secara otomatis setelah melalui tahapan- tahapan proses, yaitu Preprocessing data, pembobotan kata, membuat model untuk klasifikasi otomatis dan dibuatnya data training untuk melatih klasifikasi pada data testing. Tahapan proses tersebut dapat berjalan dengan baik dan mengklasifikasikan data dengan parameter positif dan negatif. Setelah dilakukan 3 kali pengujian didapatkan hasil akurasi 80%, 80%, dan 85%. Didapatkan hasil akurasi paling tinggi pada pengujian terakhir yakni sebesar 85%. Dengan pengujian menggunakan 3 dataset yang memiliki jumlah data yang berbeda, dan setelah mendapatkan hasil tingkat akurasi dari proses analisis sentimen dapat disimpulkan bahwa jumlah dataset dalam pengujian sangat berpengaruh terhadap tingkat akurasi Algoritma Naïve Bayes Classifier. Hal ini ditunjukan oleh hasil tingkat akurasi pada pengujian ketiga dengan 3000 dataset mendapatkan nilai akurasi 85%, lebih besar daripada pengujian pertama dengan 1000 dataset yang hanya memiliki akurasi sebesar 80%.

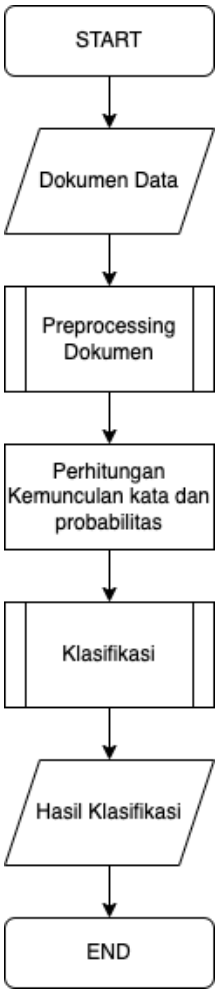
II.3.3 Sentiment Analysis Menggunakan Naïve Bayes Classifier pada Tweet Tentang Zakat

Adhyaksa H (2020), hasil klasifikasi sentiment dari 50 tweet data uji menggunakan algoritma naïve bayes dengan seleksi fitur Term-Frequency serta metode lexicon Based, didapatkan jumlah sentiment positif yang lebih dominan dibandingkan sentiment negative maupun netral dikarenakan pada pengujian dengan metode lexicon based terdapat lebih banyak tweet yang mengandung kata dalam kamus lexicon positif disbanding kata dalam kamus lexicon negative. Selanjutnya, pada pengujian dengan masing-masing seleksi fitur, sentiment positif lebih dominan dikarenakan tidak keseimbangan jumlah sentiment positif, negative dan netral dalam klasifikasi data latih menggunakan metode lexicon based dimana sentiment positif lebih besar sehingga system lebih condong dalam mengklasifikasi sentiment positif.

BAB III METODOLOGI PENELITIAN DAN PROSES BISNIS

III.1 Metode Penelitian

Metode yang digunakan dapat dilihat pada gambar 3.1:



Gambar 3. 1 urutan metode penelitian

III.2 Metodologi Pengembangan Sistem

III.2.1 Scrapping

Metode pengumpulan data melakukan scrapping yang dibantu dengan library python bernama snscrape, library ini bisa mendapatkan data users, user profiles, hasjtags, searches, tweets, list posts, communities and trend. Setelah melakukan scrapping dengan query search ” Pilpres 2024 since:2022-01-01 until:2022-12-31” penulis mendapatkan data sebanyak 93.882 data.

III.2.2 Preprocessing

Data hasil scrapping sebanyak 93.882 dibersihkan untuk mengurangi bias dan meningkatkan akurasi dari machine learning yang akan dibuat oleh penulis. Tahap preprocessing adalah tahap krusial dikarenakan hasil dari machine learning jangan sampai overfitting dan underfitting yang disebabkan oleh banyaknya bias pada saat melatih mesin pada komputer

III.2.3 Modelling

Data dari preprocessing dipisahkan menjadi data train dan test dimana perbandingan train dan test dilarang 50:50, paling tinggi ialah 60:40 dikarenakan butuh banyak Latihan pada computer untuk melakukan pengambilan keputusan. Apabila data yang dilatih dan diuji seimbang maka data akan underfitting jauh dari nilai akurasi yang diinginkan penulis. Modelling yang digunakan yaitu menggunakan metode naïve bayes classifier, yang contoh perhitungannya akan ditampilkan dibawah:

$$P(H|e) = \frac{P(e|H)P(H)}{P(e)}$$

Dari data latihan ditemukan probabilitas seperti ini:

P(pos) = 2/4

P(net) = 1/4

P(neg) = 1/4

Kemudian mencari likelihood pada setiap kata yang dicari:

$$P(H|e) = \frac{count(e,H) + 1}{count(H) + |v|}$$

Pencarian likelihood positif:

P(aku|pos) = (2+1)/(10+14) = 3/24

P(suka|pos) = (1+1)/(10+14) = 2/24

P(makanan|pos) = (0+1)/(10+14) = 1/24

Pencarian likelihood netral:

P(aku|net) = (1+1)/(10+14) = 2/24

P(suka|net) = (0+1)/(10+14) = 1/24

P(makanan|net) = (0+1)/(10+14) = 1/24

Pencarian likelihood negatif:

P(aku|neg) = (1+1)/(10+14) = 2/24

P(suka|neg) = (0+1)/(10+14) = 1/24

P(makanan|neg) = (0+1)/(10+14) = 1/24

Pencarian naïve bayes dari kalimat “aku suka makanan”:

P(pos|test) = 2/4*3/24*2/24*1/24 = 0.00021701388

P(net|test) = 1/4*2/24*1/24*1/24 = 0.00003616898

P(neg|test) = 1/4*2/24*1/24*1/24 = 0.00003616898

Dihasilkan bahwa **P(pos|test)** > P(net|test) dan P(neg|test)

III.2.4 Perangkat keras

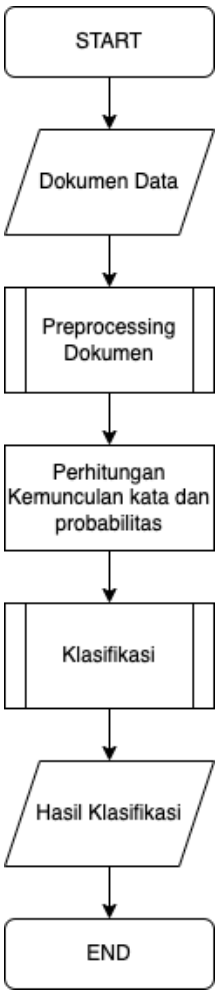
Perangkat keras yang digunakan untuk penelitian ialah:

Tabel 3. 1 spesifikasi perangkat keras

No.	Perangkat Keras	Spesifikasi
1.	Device	Macbook Pro 2020
2.	Processor	1,4 GHz Quad-Core Intel Core i5
3.	Monitor	13 inch
4.	VGA	Intel Iris Plus Graphics 645 1536 MB
5.	RAM	8 GB

III.3 Proses Bisnis

III.3.1 Proses Bisnis Sistem Berjalan



Gambar 3. 2 proses bisnis pada penelitian

Penulis memulai dengan melakukan scrapping data menggunakan library snscrape yang mana memberi query ”Pilpres 2024 since:2022-01-01 until:2022-12-31” yang berarti tweet yang ada kalimat “pilpres 2024” sejak Januari hingga akhir Desember, setelah scrapping penulis melakukan preprocessing dokumen dengan tujuan untuk membersihkan dataset yang akan digunakan sebagai bahan latihan, perhitungan dilakukan dengan menggunakan library scikit-learn sebagai library machine learning pada python