

BAB II

LANDASAN TEORI

Pada bab ini berisikan pembahasan tentang teori dasar dan definisi yang melandasi penelitian yang penulis lakukan serta tentang penelitian terkait.

2.1 Tinjauan Pustaka

2.1.1 Media Online, Crawling dan Sentimen Analisis

1. Media Online

Media Online adalah sebuah media baru dengan menggabungkan fungsi-fungsi dari teknologi internet dengan media tradisional. Fungsi Media online adalah sebagai *Surveillance*, yaitu memberikan informasi kepada khalayaknya. Tetapi bukan itu saja, media online juga mencakup semua fungsi komunikasi massa yang lain. Khalayak memiliki sifat-sifat sebagaimana yang ada pada konsep massa. Jadi khalayak media massa mempunyai sifat dan karakteristik, yaitu khalayak massa terdiri dari jumlah yang besar, ada di berbagai tempat, tidak interaktif kecuali dengan bantuan komunikasi telepon, terdiri dari lapisan masyarakat yang sangat heterogen, tidak terorganisir dan bergerak sendiri. [3]

2. Crawling

Crawling adalah proses pengambilan data dari web . *Crawler* adalah perangkat lunak atau skrip terprogram yang meramban World Wide Web dengan cara yang sistematis dan otomatis. Struktur WWW adalah struktur grafis, mis., Tautan yang disajikan dalam halaman web dapat digunakan untuk membuka halaman web lain.[4]

3. Sentimen Analisis

Sentiment analysis atau opinion mining merupakan topik riset yang penting dan sedang marak dilakukan saat ini. Opinion mining merupakan cabang penelitian dari text mining. Fokus dari opinion mining adalah melakukan analisis opini dari suatu dokumen teks.

Terdapat tiga buah subproses dari opinion mining yaitu, document subjectivity, opinion orientation dan target detection. Dalam dunia bisnis, opinion mining banyak digunakan untuk menganalisis secara otomatis opini pelanggan tentang produk dan pelayanannya.

Menurut Pozzi[5] Analisis sentimen merupakan salah satu riset yang kompleks. Adapun karakteristik dari Analisis Sentimen adalah sebagai berikut.

1. Kategorisasi Sentimen yaitu membedakan antara kalimat subjektif dan kalimat objektif.
2. Tingkat analisis. Tingkat analisis dibagi menjadi 3 tingkatan yaitu Message level, Sentence Level dan Entity and Aspect Level.
3. Pendapat yang membandingkan sesuatu serta pendapat yang hanya sekedar pendapat. Artinya setiap orang dapat memberikan pendapat dengan membandingkan satu hal dengan hal lain atau hanya memberikan pendapatnya saja.
4. Pendapat eksplisit dan pendapat implisit. Pendapat yang diungkapkan secara jujur, tegas dan jelas atau pendapat yang diungkapkan secara tidak jelas.

4. Polaritas Sentiment

Setelah menentukan skor sentimen dari setiap kata maka tahap selanjutnya adalah menentukan nilai sentimen keseluruhan dari setiap kalimat. Nilai sentimen akan di tentukan dengan menggunakan rumus sebagai berikut:

$$S_{positive} = \sum_{i \in T}^n positive\ score_i$$

$$S_{negative} = \sum_{i \in T}^n negative\ score_i$$

Kedua persamaan di atas merupakan rumus untuk menghitung orientasi semantic dari suatu kalimat. Skor positif dan skor negatif dari masing-masing suku kata pada satu kalimat akan dijumlahkan secara terpisah, selanjutnya untuk mengetahui orientasi semantik dari suatu kalimat apakah bernilai negatif atau bernilai positif menggunakan rumus berikut:

$$Sentence_{sentiment} \begin{cases} positive\ if\ S_{positive} > S_{negative} \\ neutral\ if\ S_{positive} = S_{negative} \\ negative\ if\ S_{positive} < S_{negative} \end{cases}$$

Orientasi semantik dari suatu kalimat didapat dengan membandingkan antara skor negatif dan skor positif. Hasilnya bernilai positif jika skor positif lebih besar dari skor negatif, sedangkan hasilnya bernilai negatif jika skor positif lebih kecil skor negatif, jika skor positif sama dengan skor positif maka hasilnya netral[6]

Kata kunci diekstrak dari kalimat menggunakan token yang dicocokkan dengan kamus untuk mendapatkan kata kunci positif dan negatif. Prosesnya bisa dilihat sebagai berikut.

- Ekstraksi kata kunci positif dan kata kunci negative Token-token dicocokkan dengan kata kunci yang ada dalam kamus yang berisikan kata kunci positif ataupun kata kunci negatif.

- Ekstraksi kata sifat yang dicurigai sebagai kata opini Kata sifat yang dicurigai sebagai kata opini diekstrak menjadi kata opini kemudian ditambahkan dalam kamus opini.

Contohnya:

“banyak” adalah kata sifat, akan tetapi juga sebagai kata opini dan terdapat dalam kamus positif maka akan menjadi kata opini positif.

- Evaluasi negasi

Dengan menggunakan kata negasi, kata yang dideteksi positif dapat berubah menjadi negative maupun sebaliknya. Contohnya “korupsi” adalah kata kunci negatif, akan tetapi jika sebelumnya terdapat kata “tidak” dan menghasilkan token “tidak korupsi” maka nilainya akan berubah menjadi positif. Kata-kata negasi tersebut, dapat dilihat dari kamus yang berisikan kata negasi.

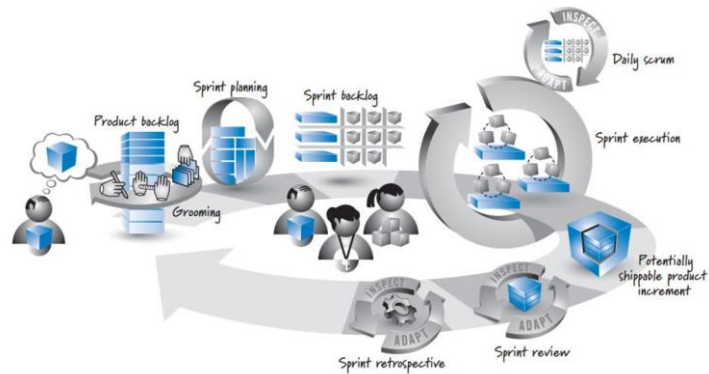
2.1.2 Metodologi Pengembangan Scrum

Scrum (kb): Sebuah kerangka kerja dimana orang-orang dapat mengatasi masalah kompleks adaptif, dimana pada saat bersamaan mereka juga menghantarkan produk dengan nilai setinggi mungkin secara produktif dan kreatif.

Scrum adalah sebuah kerangka kerja untuk mengembangkan, menghantarkan dan mengelola produk yang kompleks. Panduan ini berisi definisi Scrum yang terdiri dari peran-peran, acara-acara, artefak-artefak, dan aturan-aturan yang menyatukan kesemuanya. Ken Schwaber dan Jeff Sutherland mengembangkan Scrum; Panduan Scrum ditulis dan disebarluaskan oleh mereka. Mereka berdua yang tetap mempertahankan Panduan Scrum.

Kerangka kerja Scrum terdiri dari Scrum Team dan peran-peran, acara-acara, artefak-artefak dan aturan-aturan terkait. Setiap komponen di dalam kerangka kerja

ini memiliki tujuan tertentu dan sangat penting bagi keberhasilan penggunaan Scrum.[7]



Gambar 1 : Gambar Flow Scrum

1. Product Backlog

Product Backlog adalah daftar terurut semua hal yang telah diketahui hingga saat ini harus ada di dalam produk. Product Backlog adalah satu-satunya sumber kebutuhan untuk semua perubahan yang perlu diberlakukan terhadap produk. Product Owner bertanggung jawab terhadap Product Backlog, termasuk isi, ketersediaan dan urutannya.

Sebuah Product Backlog tidak pernah tuntas. Di awal pengembangan produk, Product Backlog berisi daftar kebutuhan yang diketahui dan dipahami saat ini. Product Backlog berevolusi seiring dengan perkembangan produk dan lingkungan dimana produk tersebut digunakan. Product Backlog bersifat dinamis dan berubah terus secara konstan agar produk menjadi layak, kompetitif dan bermanfaat. Selama produk masih ada, Product Backlog juga akan selalu ada. [7]

2. Sprint

Jantung dari Scrum adalah Sprint, yaitu sebuah batasan waktu dengan durasi satu bulan atau kurang, dimana terdapat proses pembuatan Increment yang “Selesai”, dapat digunakan dan berpotensi untuk dirilis. Sprint memiliki durasi yang

konsisten sepanjang daur hidup pengembangan produk. Sprint yang baru langsung dimulai setelah Sprint sebelumnya selesai. Sprint berisi dan terdiri dari Sprint Planning, Daily Scrum, pengembangan produk, Sprint Review dan Sprint Retrospective.

Pada saat Sprint berjalan:

- Tidak boleh ada perubahan yang dapat mengancam Sprint Goal;
- Tingkat kualitas tidak boleh menurun; dan,
- Ruang lingkup dapat diklarifikasi dan dinegosiasi ulang antara Product Owner dan Development Team setiap kali adanya hal baru yang mereka pelajari;

Setiap Sprint bisa dianggap sebagai sebuah proyek dengan durasi tidak lebih dari satu bulan yang dijalankan untuk mencapai sebuah tujuan. Setiap Sprint memiliki tujuan mengenai apa yang harus dibangun, sebuah rancangan dan perencanaan fleksibel yang memandu pembangunan tersebut, daftar pekerjaan, dan hasil dari Increment. [7]

3. Sprint Goals

Sprint Goal adalah sebuah objektif untuk Sprint yang dapat dicapai lewat pengimplementasian Product Backlog. Sprint Goal merupakan panduan bagi Development Team untuk menjawab pertanyaan mengapa mereka mengembangkan Increment. Sprint Goal dibuat pada pertemuan Sprint Planning. Sprint memberikan ruang fleksibilitas mengenai fungsionalitas yang akan diimplementasikan di dalam Sprint. Product Backlog item terpilih merupakan satu fungsi yang selaras yang bisa jadi Sprint Goal. Sprint Goal bisa saja menghubungkan pekerjaan yang tidak memiliki keterkaitan agar Development Team tidak bekerja dari instruksi kerja yang berbeda-beda.[7]

4. Sprint Review

Sprint Review diselenggarakan di akhir Sprint untuk menginspeksi Increment dan mengadaptasi Product Backlog bila diperlukan. Pada saat Sprint Review, Scrum

Team dan pemegang kepentingan berkolaborasi untuk meninjau apa yang sudah diselesaikan di Sprint. Berdasarkan hasil tinjauan tersebut dan perubahan terhadap Product Backlog di dalam Sprint, hadirin berkolaborasi untuk menentukan pekerjaan selanjutnya yang dapat dilakukan untuk mengoptimalkan nilai bisnis. Ini adalah pertemuan informal, bukan pertemuan laporan status, dan presentasi Increment dilakukan guna mendapatkan umpan balik dan mengembangkan kemampuan kolaborasi.

Pertemuan ini paling lama diselenggarakan selama empat jam untuk Sprint berdurasi satu bulan. Untuk Sprint yang lebih singkat, durasi acara ini biasanya lebih singkat. Scrum Master memastikan acara ini terselenggara dan hadirin memahami tujuannya. Scrum Master mengedukasi setiap hadirin untuk menjaganya di dalam batasan waktu. [7]

5. Sprint Retrospective

Sprint Retrospective adalah sebuah kesempatan bagi Scrum Team untuk menginspeksi dirinya sendiri dan membuat perencanaan mengenai peningkatan yang akan dilakukan di Sprint berikutnya.

Sprint Retrospective terselenggara setelah Sprint Review dan sebelum Sprint Planning berikutnya. Acara ini diselenggarakan paling lama tiga jam untuk Sprint yang berdurasi satu bulan. Untuk Sprint yang lebih singkat, durasi acara ini biasanya lebih singkat. Scrum Master memastikan acara ini terselenggara dan setiap peserta memahami tujuannya. [7]

6. Sprint Planning

Pekerjaan yang akan dikerjakan di Sprint direncanakan pada saat Sprint Planning. Perencanaan ini dilakukan secara kolaboratif oleh seluruh anggota Scrum Team.

Sprint Planning memiliki batasan waktu maksimal delapan jam untuk Sprint yang berdurasi satu bulan. Untuk Sprint yang lebih singkat, acara ini biasanya lebih singkat. Scrum Master memastikan acara ini diselenggarakan dan peserta

memahami tujuannya. Scrum Master mendukung Scrum Team untuk menjaganya di dalam batasan waktu. [7]

7. Scrum Master

Scrum Master bertanggung jawab untuk mengenalkan dan menyokong penggunaan Scrum sebagaimana dijelaskan di dalam Panduan Scrum ini. Scrum Master melakukan ini dengan membantu orang-orang agar dapat memahami teori, praktik-praktik, aturan-aturan dan tata nilai Scrum.

Scrum Master adalah pemimpin yang melayani Scrum Team. Scrum Master membantu orang-orang di luar Scrum Team untuk dapat memahami interaksi mana yang bermanfaat dan tidak bermanfaat. Scrum Master membantu orang-orang untuk mengubah interaksi ini guna memaksimalkan nilai bisnis yang dihasilkan oleh Scrum Team. [7]

2.1.3 Infrastruktur Pengembangan

1. Big Data

Big Data adalah kombinasi teknologi yang bisa memanajemen data yang bervariasi dalam jumlah besar, kecepatan yang tepat, dan memiliki ketepatan saat melakukan analisis dan reaksi. Tiga karakteristik yang dimiliki Big Data, yaitu volume, velocity, dan variety. [8]

2. Elastic Search

Elasticsearch adalah mesin pencarian teks dan analitik yang bersifat opensource yang sangat scalable serta memungkinkan penggunaannya untuk menyimpan, mencari, dan menganalisis data dengan volume yang besar dengan cepat dan mendekati real time. Elasticsearch umumnya digunakan sebagai mesin/teknologi yang mampu mengoptimasi sebuah aplikasi yang memiliki fitur pencarian.

Penelitian Bhadane, et. Al [9] mengoptimalkan perangkat lunak yang mereka bangun dengan menambahkan Elasticsearch sehingga mampu memudahkan pengguna dalam mengambil daftar obat-obatan berdasarkan kata kunci tertentu. Selain itu, Elasticsearch digunakan untuk meningkatkan query pada database konvensional sehingga membuatnya lebih cepat sebelum mengirimnya ke server. Hal ini mampu mengurangi latensi waktu yang digunakan ketika melakukan pencarian.

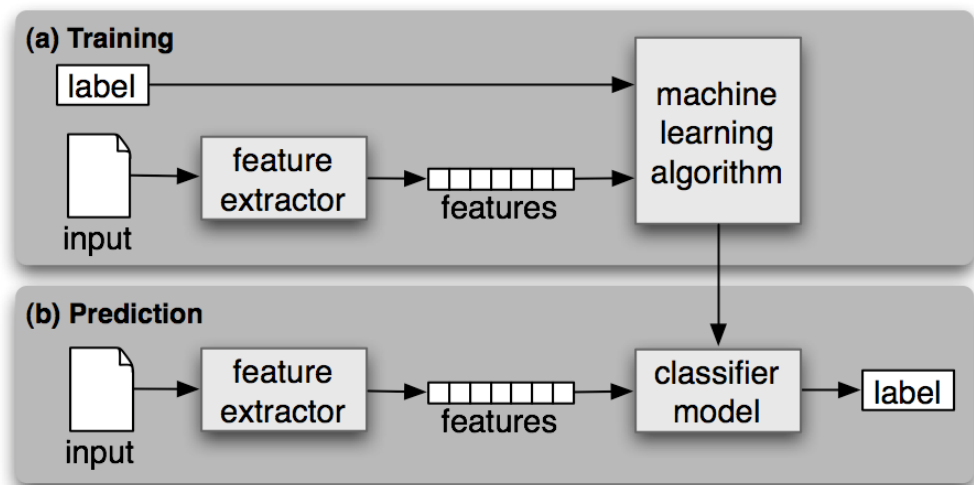
3. Python

Python merupakan bahasa pemrograman tingkat tinggi. Hal ini disebabkan karena kode yang dituliskan akan di compile menjadi byte code dan dieksekusi sehingga Python cocok digunakan untuk scripting language, aplikasi web dan lain sebagainya. Hal lain yang menjadikan bahasa ini menjadi bahasa pemrograman tingkat tinggi adalah Python dapat di extend kedalam bahasa C dan C++ serta bahasa pemrograman ini memiliki struktur konstruksi yang kuat (blok kode, fungsi, class, modules, dan packages) dan serta konsisten menggunakan konsep Object Oriented Programming (OOP). [10]

5. NLTK

NLTK merupakan salah satu tools pengolahan bahasa natural yang berjalan pada bahasa pemrograman Python. NLTK menawarkan tampilan antar muka yang mudah dipahami dan menyediakan lebih dari 50 data serta kamus data yang dapat digunakan. Beberapa kamus data yang dapat digunakan antara lain WordNet, TextProcessing serta untuk proses klasifikasi yaitu tokenization, stemming, tagging, parsing dan juga semantic reasoning.

NLTK juga mempunyai tools untuk melakukan klasifikasi document menggunakan naïve bayes yang bisa diimplementasikan untuk menentukan sentiment negatif dan positif. [11]



Gambar 2 : Implementasi Naïve Bayes Classifier NLTK

6. Sastrawi stemmer

Sastrawi stemmer merupakan sebuah library stemmer sederhana yang didesain untuk dapat digunakan secara mudah. Sastrawi stemmer berfungsi untuk mengubah kata berimbuhan menjadi kata dasar. Aturan-aturan bahasa diterapkan untuk menanggalkan imbuhan-imbuhan itu.

Ada beberapa algoritma Stemming Bahasa Indonesia, semuanya menghadapi tantangan yang serupa. Sastrawi stemmer menerapkan algoritma yang berbasis Nazief dan Adriani, kemudian ditingkatkan oleh Algoritma CS (Confix Stripping), kemudian ditingkatkan lagi oleh algoritma ECS (Enhanced Confix Stripping), lalu ditingkatkan lagi oleh Modified ECS. [12]

Stemming adalah proses yang menyediakan pemetaan variasi kata morfologis yang berbeda ke dalam kata dasar / umum mereka. Proses ini juga dikenal sebagai penggabungan. Berdasarkan anggapan bahwa istilah-istilah yang memiliki stem yang sama biasanya akan memiliki arti yang sama, stem proses ini banyak digunakan dalam Pengambilan Informasi sebagai cara untuk meningkatkan kinerja pengambilan. selain kemampuannya untuk meningkatkan

kinerja pengambilan, proses stemming, yang dilakukan pada saat pengindeksan, juga akan mengurangi ukuran file indeks. [13]

7. Sastrawi Stopwords Remover

Sastrawi stopwords remover merupakan sebuah library remover words sederhana yang digunakan untuk menghapus kata umum (common words) yang biasanya muncul dalam jumlah besar dan dianggap tidak memiliki makna. Stop words umumnya dimanfaatkan dalam task information retrieval, termasuk oleh Google. Contoh stop words untuk bahasa Inggris diantaranya “of”, “the”. Sedangkan untuk bahasa Indonesia diantaranya “yang”, “di”, “ke”. Kamus stopwords dari sastrawi dapat ditambahkan oleh data eksternal secara manual.[14]

8. Confusion Matrix

Confusion matrix adalah suatu metode yang biasanya digunakan untuk melakukan perhitungan akurasi pada konsep data mining. Confusion matrix digambarkan dengan tabel yang menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan.[16]

Table 1 : Confusion Matrix

Correct Classification	Classified as	
	Predicted “+”	Predicted “-”
Actual “+”	True Positives	False Negative
Actual “-”	False Positives	True Negatives

Berdasarkan tabel Confusion Matrix diatas:

- True Positives (TP) adalah jumlah record data positif yang diklasifikasikan sebagai nilai positif
- False Positives (FP) adalah jumlah record data negatif yang diklasifikasikan sebagai nilai positif

- c. False Negatives (FN) adalah jumlah record data positif yang diklasifikasikan sebagai nilai positif
- d. True Negatives (TN) adalah jumlah record data negatif yang diklasifikasikan sebagai nilai negative

Nilai yang dihasilkan melalui metode Confusion Matrix adalah berupa evaluasi sebagai berikut :

- a. Accuracy, presentase jumlah record data yang diklasifikasikan (prediksi) secara benar oleh algoritma.

Rumus : $(TP + TN) / \text{Total data} = \text{Accuracy}$.

- b. Misclassification (Error) Rate, presentase jumlah record data yang diklasifikasikan (prediksi secara salah oleh algoritma).

Rumus : $(FP + FN) / \text{Total data} = \text{Misclassification Rate}$.

2.1.4 Algoritma Naïve Bayes

Naive Bayes merupakan sebuah model yang dapat bekerja dengan baik pada proses pembagian kategori untuk teks. Naive bayes melakukan klasifikasi dengan menggunakan dua proses yaitu proses training dan proses testing.

Algoritma Naive Bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naive dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi Naive Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya.

Persamaan dari teorema Bayes adalah :

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Keterangan :

X : Data dengan class yang belum diketahui

H : Hipotesis data X merupakan suatu class spesifik

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posteriori probability)

$P(H)$: Probabilitas hipotesis H (prior probability)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H $P(X)$: Probabilitas

Untuk menjelaskan teorema Naive Bayes, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema bayes di atas disesuaikan sebagai berikut :

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)}$$

Dimana Variabel C merepresentasikan kelas, sementara variabel $F_1 \dots F_n$ merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (Posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut prior), dikali dengan peluang kemunculan karakteristik – karakteristik sampel pada kelas C (disebut juga likelihood), dibagi dengan peluang kemunculan karakteristik – karakteristik sampel secara global (disebut juga evidence). Karena itu, rumus diatas dapat pula ditulis secara sederhana sebagai berikut :

$$\text{Posterior} = \frac{\text{Prior} \times \text{likelihood}}{\text{evidence}}$$

Nilai Evidence selalu tetap untuk setiap kelas pada satu

$$P(X_i = x_i | Y = y_j) = \frac{1}{2\pi\sigma_{ij}} e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}}$$

sampel. Untuk klasifikasi dengan data kontinyu digunakan rumus Densitas Gauss :

Di mana :

P : Peluang

X_i : Atribut ke i

x_i : Nilai atribut ke i

Y : Kelas yang dicari

y_j : Sub kelas Y yang dicari

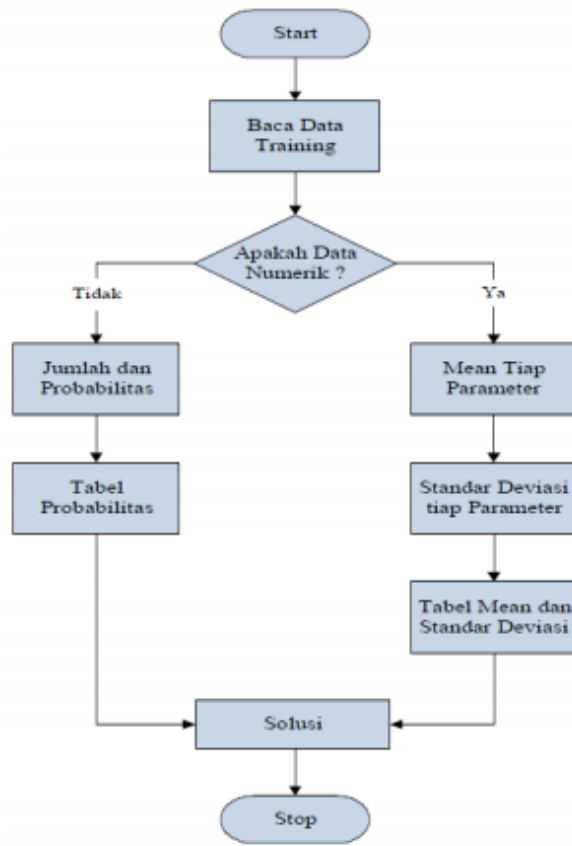
μ_{ij} : Mean, menyatakan rata – rata dari seluruh atribut

σ_{ij} : Deviasi standar, menyatakan varian dari seluruh atribut

Adapun alur dari metode Naive Bayes adalah sebagai berikut :



STT - NF



Gambar 3: Alur Metode Naive Bayes

Penjelasan dari gambar diatas:

1. Baca data training
2. Hitung Jumlah dan probabilitas, namun apabila data numerik maka:
 - a. Cari nilai mean dan standar deviasi dari masing – masing parameter yang merupakan data numerik.
 - b. Cari nilai probabilitik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
3. Mendapatkan nilai dalam tabel mean, standart deviasi dan probabilitas.
4. Solusi kemudian dihasilkan. [15]

2.1.5 Blackbox Testing

Menurut Rex Black, black box testing adalah suatu metode pengujian dimana tester hanya fokus pada apa yang seharusnya dilakukan oleh sistem [17]. Sebuah tes dapat dikatakan berhasil ketika sebuah sistem dapat memproses data dan hasil yang ada sesuai dengan apa yang diharapkan. Ketika menggunakan metode black box, tester tidak perlu mengetahui bagaimana struktur dan desain data yang ada di dalam sistem. Mereka hanya melihat apakah sistem terjadi bugs atau tidak. [18]

Uji coba blackbox berusaha untuk menemukan kesalahan dalam beberapa kategori, diantaranya:

1. Fungsi-fungsi yang salah atau hilang.
2. Kesalahan interface.
3. Kesalahan dalam struktur data atau akses database eksternal.
4. Kesalahan performa.
5. kesalahan inisialisasi dan terminasi

STT - NF

2.2 Penelitian Terkait

Table 2 : Penelitian Terkait

Nama, Tahun dan Institusi	Judul	Sumber data crawling	Tools	Metodologi	Hasil/ Kesimpulan
Sigit Suryono, Ema Utami, Taufiq Luthfi, 2017	KLASIFIKASI SENTIMEN PADA TWITTER DENGAN NAIVE BAYES CLASSIFIER	Twitter	Python	Naive Bayes	Hasil dari skenario uji coba : Berdasarkan hasil klasifikasi besar persentase untuk ketiga kelas, ditempat pertama yaitu dengan kelas sentimen neutral dengan 52%, kedua sentimen positif dengan 28% dan ketiga sentimen negatif dengan 20%. Berdasarkan hasil persentase kelas sentimen, sentimen neutral merupakan sentimen yang paling banyak apabila dikaitkan dengan topik Presiden Joko Widodo dan Pemerintahannya.
Jaka Eka Sembodo, Erwin Budi Setiawan, ZK Abdurahman Baizal, 2016 (Computational Science, School of Computing, Telkom University)	Data Crawling Otomatis pada Twitter	Twitter	Php	-	Implementasi crawling data di twitter secara otomatis dalam penelitian ini menggunakan Bahasa pemograman PHP dan basisdata MySQL. Interface (API) twitter.
Bustami, 2014 (Teknik Informatika, Universitas Malikussaleh Reuleut, Aceh Utara, Aceh-Indonesia)	PENERAPAN ALGORITMA NAIVE BAYES UNTUK MENGLASIFIKASI DATA NASABAH ASURANSI	Nasabah Asuransi	Visual Basic	Naive Bayes	Sistem klasifikasi data nasabah ini digunakan untuk menampilkan informasi klasifikasi lancar, kurang lancar atau tidak lancarnya calon nasabah dalam membayar premi asuransi dengan menggunakan algoritma Naive Bayes.
Muhamad Fadri Wijaya, 2019 (Teknik Informatika, STT Nurul Fikri, Depok, Jawa Barat, Jawa-Indonesia)	Analisis Sentimen Hasil Crawling pada Media Online Menggunakan Elasticsearch dan Algoritma Naive Bayes Classifier.	Media online (kompas, detik, tribunnews, kumparan)	Python, Elasticsearch	Naive Bayes	-

2.3 Posisi Penelitian

Table 3: Posisi Penelitian

No.	Algoritma Naïve Bayes	Crawling	Media Online / Sosial Media	Python	ElasticSearch
1	Bustami (Teknik Informatika, Universitas Malikussaleh Reuleut, Aceh Utara, Aceh-Indonesia) PENERAPAN ALGORITMA NAIVE BAYES UNTUK MENGLASIFIKASI DATA NASABAH ASURANSI				
2		Jaka Eka Sembodo , Erwin Budi Setiawan , ZK Abdurahman Baizal (Computational Science, School of Computing, Telkom University) Data Crawling Otomatis pada Twitter			
3	Sigit Suryono, Em a Utami, Em ha Taufiq Luthfi (Universitas Amikom Yogyakarta) KLASIFIKASI SENTIMEN PADA TWITTER DENGAN NAIVE BAYES CLASSIFIER				
4	Muhamad Fadri Wijaya (Teknik Informatika, STT Nurul Fikri, Depok, Jawa Barat, Jawa-Indonesia) Analisis Sentimen Hasil Crawling pada Media Online Menggunakan Elasticsearch dan Algoritma Naive Bayes Classifier.				

STT - NF