UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

BAB II

LANDASAN TEORI

2.1 Shopee

Shopee adalah *mobile-platform* pertama di Asia Tenggara (Indonesia, Filipina, Malaysia, Singapura, Thailand, Vietnam) dan Taiwan yang menawarkan transaksi jual beli online yang menyenangkan, gratis, dan terpercaya via ponsel. Kamu dapat mendaftarkan produk jualan dan berbelanja berbagai penawaran menarik dengan harga termurah dan gratis ongkir ke seluruh Indonesia.

Shopee mulai masuk ke pasar Indonesia pada akhir bulan Mei 2015 dan Shopee baru mulai beroperasi pada akhir Juni 2015 di Indonesia. Shopee merupakan anak perusahaan dari Garena yang berbasis di Singapura. Shopee telah hadir di beberapa negara di kawasan Asia Tenggara seperti Singapura, Malaysia, Vietnam, Thailand, Filipina, dan Indonesia.Shopee Indonesia beralamat di Wisma 77 Tower 2, Jalan Letjen. S. Parman, Palmerah, Daerah Khusus Ibukota Jakarta 11410, Indonesia.

Sasaran pengguna Shopee adalah kalangan muda yang saat ini terbiasa melakukan kegiatan dengan bantuan *gadget* termasuk kegiatan berbelanja. untuk itu Shopee hadir dalam bentuk aplikasi *mobile* guna untuk menunjang kegiatan berbelanja yang mudah dan cepat. Kategori produk yang ditawarkan Shopee lebih mengarah pada produk *fashion* dan perlengkapan rumah tangga.

Jenis Produk yang ditawarkan:

1. Pakaian Pria dan Wanita (*Fashion*)
2. *Handphone* dan Aksesoris
3. Komputer dan Aksesoris
4. Perawatan dan Kesehatan
5. Perlengkapan Rumah
6. Sepatu Pria dan Wanita
7. Tas Pria dan Wanita
8. Elektronik
9. Fotografi

10. Makanan dan Minuman Dan masih banyak lagi.

Saat ini Shopee sudah dapat mencakup wilayah diseluruh Indonesia bahkan dikota kecil. Dan sudah banyak Penjual yang menawarkan Produk nya pada aplikasi Shopee dan banyak juga Konsumen memilih Shopee sebagai tempat Belanja Online.

Adapun cara Praktis pada Shopee melakukan Pembayaran, yaitu:

1. Kartu Kredit / Debit *Online*
2. *Indomart*
3. *Transfer* Bank
4. Pembayaran Kredivo

Di Shopee untuk pembayaran via *Transfer* Bank sangat mudah karna ada aplikasi Pengecekan Otomatis, pembeli tidak perlu upload bukti transfer. Sedangkan Pembayaran Kredivo adalah kredit instan yang memberikan pembeli kemudahan untuk bayar dalam 30 hari tanpa bunga atau dengan fasilitas cicilan 3, 6 & 12 bulan tanpa memerlukan kartu kredit. Omset Perusahaan tidak diketahui, tetapi adapun Omset beberapa Penjual di Shopee. Salah satunya ada Jordan Atik Bawazier, pemilik toko online @PakaianOnline di Shopee, Jordan dapat mencapai Rp3,5 miliar setiap bulan.

Sebagai sebuah media sosial yang digunakan banyak khayalak ramai, tentunya Shopee memiliki beberapa kelebihan dan kekurangan. Berikut kelebihan dan kekurangan Shopee:

1. Kelebihan shopee

a. Banyaknya kategori yang disediakan
b. Respon *website* yang cepat
c. Tampilannya simpel dan menarik
d. Tersedia penjelasan spesifik barang.
e. Adanya pengiriman gratis tanpa ongkos kirim
f. Adanya pilihan gratis ongkir

UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

2. Kelemahan shopee

a. Ongkos kirim rusak di tanggung pembeli
b. Untuk promo ongkos kirim gratis syarat yang harus dipenuhi terkesan sulit dan merepotkan
c. Terkadang situs sulit di akses atau bahkan lambat pada jam-jam tertentu
d. Tidak adanya layanan belanja dengan cicilan (pembeli harus membayar barang secara *cash*)
e. Terkadang jika membeli dua barang di toko yang sama ada salah satu barang yang tidak sampai pada pelanggan.

2.2 *Google play*

Google play adalah layanan konten digital milik *Google* yang melingkupi toko daring untuk produk-produk seperti musik/lagu, buku, aplikasi, permainan ataupun pemutar media yang berbasis awan. Layanan *Google Play* ini dapat diakses baik melalui web, aplikasi android (*Play Store*), dan *Google* TV. *Google Play* mulai dikenalkan pada bulan Maret 2012 sebagai pengganti dari Android Market dan Layanan Musik *Google* (*Google Play, 2017*).

2.3 *Text Mining*

Text mining merupakan bagian penerapan dari data mining yang digunakan untuk mencari pola dalam teks. Berbeda dengan data mining yang menggunakan data biasa, dalam *text mining* ini data yang digunakan adalah data dalam bentuk dokumen atau teks (Feizar, 2014).

Text mining mempunyai kesamaan dengan data mining. *Text mining* mempunyai tujuan yaitu untuk memperoleh informasi dan pengetahuan yang berguna dari sekumpulan data yang sangat besar. Data tersebut bisa berbentuk sebuah database. Namun keduanya memiliki perbedaan pada jenis data. *Data mining* memiliki masukan data dari data yang sudah terstruktur sedangkan *text mining* umumnya menggunakan data yang tidak terstruktur atau minimal semi terstruktur. Hal ini menyebabkan adanya tantangan tambahan pada *text mining* yaitu struktur teks yang kompleks dan tidak lengkap, arti yang tidak jelas dan

tidak standar, dan bahasa yang berbeda ditambah translasi yang tidak akurat (Sussolaikah & Alwi, n.d.2016).

2.3.1 *Text Preprocessing*

Text preprocessing merupakan tahapan dari proses awal terhadap teks untuk mempersiapkan teks menjadi data yang nantinya akan diolah lebih lanjut. Sebuah teks yang ada harus dipisahkan, hal ini dapat dilakukan dalam beberapa tingkatan yang berbeda. Suatu dokumen dapat dipecah menjadi bab, sub-bab, paragraf, kalimat dan pada akhirnya menjadi potongan kata/token. Selain itu pada tahapan ini keberadaan digit angka, huruf kapital, atau karakter-karakter yang lainnya dihilangkan dan dirubah (Larasati, 2015).

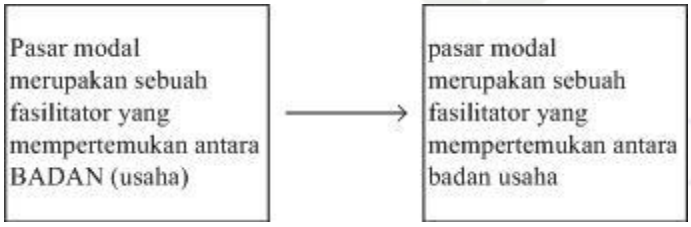
Adapun tahapan *preprocessing* berdasarkan proses yaitu (Setiawan, Fitri Astuti, & Harsa Kridalaksana, 2015):

1. *Cleaning*

   Cleaning merupakan proses untuk membersihkan dokumen dari komponen-komponen yang tidak memiliki hubungan dengan informasi yang ada pada dokumen.

2. *Case Folding*

   Case folding merupakan tahapan yang mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf ‘a’ sampai dengan ‘z’ yang diterima. Karakter selain huruf dihilangkan dan dianggap delimiter (pembatas).

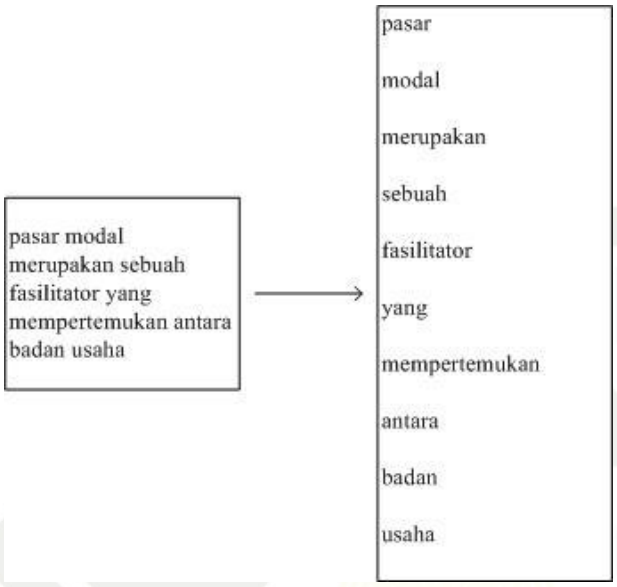


Gambar 2.1 Proses *Case Folding*

UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

3. *Tokenizing*

Tokenizing / parsing adalah tahap pemotongan string input berdasarkan tiap kata yang menyusunnya. Selain itu, spasi digunakan untuk memisahkan antar kata tersebut.



Gambar 2.2 Proses *Tokenizing*

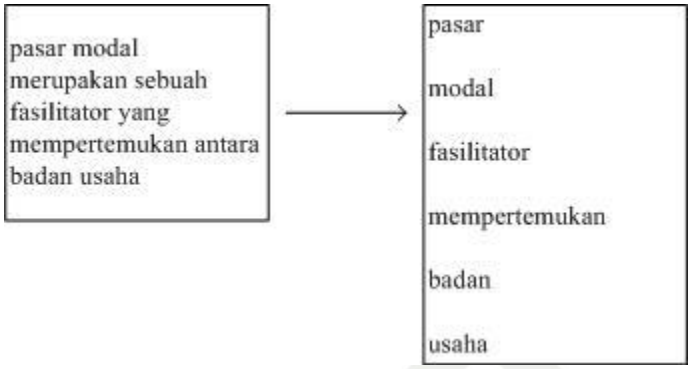
4. Normalisasi

Merupakan proses pengembalian kata-kata yang tidak baku kedalam bahasa baku dalam kamus bahasa indonesia.

5. *Filtering*

Filtering adalah tahap mengambil kata-kata penting dari hasil *tokenizing*. Proses *filtering* dapat menggunakan algoritma *stoplist* (membuang kata yang kurang penting) atau *wordlist* (menyimpan kata penting). *Stoplist* / *stopworld* adalah kata-kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-words*.

© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 2. 3 Proses *Filtering*

6. Stemming

Stemming merupakan proses yang digunakan untuk menemukan kata dasar dari sebuah kata dengan menghilangkan semua imbuhan (*affixes*) baik yang merupakan awalan (*prefixes*), sisipan (*infixex*), akhiran (*suffixes*), dan kombinasi dari awalan serta akhiran (*confixes*) pada kata turunan.

Algoritma *stemming* kata pada bahasa Indonesia dengan performa yang paling baik saat ini (memiliki jenis kesalahan *stemming* yang paling sedikit) adalah Algoritma *Enhanced Confix Stripping* (ECS) (Arifin, Mahendra, & Ciptaningtyas, 2009). Algoritma *stemming* ini diperkenalkan oleh Jelita Asian sebagai pengembangan dari algoritma Nazief dan Adriani pada tahun 1996. Algoritma ECS merupakan algoritma perbaikan dari algoritma *stemming Confix Stripping* (CS). Perbaikan yang dilakukan oleh stemming ECS adalah perbaikan beberapa aturan pada tabel acuan pemenggalan imbuhan. Selain itu, algoritma stemming ECS juga menambahkan langkah pengembalian akhiran jika terjadi penghilangan akhiran yang seharusnya tidak dilakukan (Tahitoe & Purwitasari, 2010).

[DP+[DP + [DP+]]] Kata Dasar [[+DS][+PP][+P]

UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

Keterangan :

DP (*Derivation Prefixes*) : kumpulan awalan yang diberikan langsung pada kata dasar

DS (*Derivation Suffixes*) : kumpulan akhiran yang ditambahkan langsung pada kata dasar

PP (*Possessive Pronoun*) : kata ganti kepunyaan

Particle : partikel

Algoritma *Enhanched Confix Stripping* ini mempunyai beberapa tahapan proses sebagai berikut (Prasetyo & Susanti, 2012) :

1. Kata yang hendak di *stemming* dicari terlebih dahulu pada kamus. Jika ditemukan, berarti kata tersebut adalah kata dasar, jika tidak maka langkah 2 dilakukan.
2. Cek *rule precedence*. Apabila suatu kata memiliki pasangan awalan-akhiran "be-lah", "be-an", "me-i", "di-i", "pe-i", atau "te-i" maka langkah stemming selanjutnya adalah (5, 3, 4, 6). Apabila kata tidak memiliki pasangan awalan- akhiran tersebut, langkah stemming berjalan normal (3, 4, 5, 6).
3. Hilangkan *inflectional particle* P ("-lah", "-kah", "-tah", "-pun") dan kata ganti kepunyaan atau possessive pronoun PP ("-ku", "-mu", "-nya").
4. Hilangkan *Derivation Suffixes* DS ("-i", "-kan", atau "-an").
5. Hilangkan *Derivational Prefixes* DP {"di-","ke-","se-","me-","be-","pe", "te-"}.

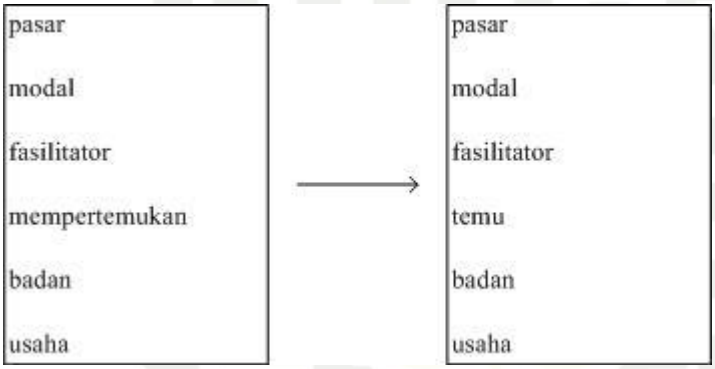
a. Identifikasikan tipe awalan dan hilangkan. Awalan ada dua tipe:
   1. Standar: "di-", "ke-", "se-" yang dapat langsung dihilangkan dari kata.
   2. Kompleks: "me-", "be-", "pe", "te-" adalah tipe-tipe awalan yang dapat bermorfologi sesuai kata dasar yang mengikutinya.

© Hak Cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

b. Cari kata yang telah dihilangkan awalannya ini di dalam kamus. Apabila tidak ditemukan, maka langkah 5 diulangi kembali. Apabila ditemukan, maka keseluruhan proses dihentikan.

6. Jika semua gagal, maka masukan kata yang diuji pada algoritma ini dianggap sebagai kata dasar.

Tahap *stemming* adalah tahap mencari root kata dari tiap kata hasil filtering. Pada tahap ini dilakukan proses pengembalian berbagai bentukan kata ke dalam suatu representasi yang sama.



Gambar 2.4 Proses *Stemming*

2.4 *Negation Handing*

Untuk meningkatkan akurasi dalam klasifikasi juga dapat menerapkan *negation handling* dalam prosesnya, dimana apabila terdapat kata yang bersifat negasi seperti tidak, bukan, nggak dan lainnya, hal ini akan membuat makna kata berikutnya menjadi berlawanan dengan makna sesungguhnya. Berikut ini beberapa langkah-langkah *negation handling* menurut Vivek Narayanan (Narayanan, 2013).

1. Inisialisasi negasi = *false*
2. For *each* kata *in* dok

    Ifnegasi = *True*

    Ubah kata menjadi = “kata_negasi” + “kata”

    If kata = “tidak” or “bukan” or “nggak” or “belum”

    Negasi = *True*

    Elsenegasi = *false*

UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

Proses *negation handling* ini yang dilakukan adalah mencari kata yang bersifat negasi, jika ditemukan kata negasi maka kata negasi tersebut digabung dengan kata selanjutnya dan selanjutnya digantikan dengan lawan kata setelah kata negasi tersebut.

2.5 Naïve Bayes Classifier

Algoritma *Naive Bayes Classifier* merupakan algoritma metode pengklasifikasian turunan dari teorema Bayes yang sederhana, berguna untuk mencari nilai probabilitas atau peluang tertinggi untuk mengklasifikasikan data testing (uji) pada kategori yang paling tepat (Feldman, 2007). Menurut Kusumadewi (2009) metode *naïve bayes classifier* adalah algoritma klasifikasi yang sangat efektif (mendapatkan hasil yang tepat) dan efisien (proses penalaran dilakukan memanfaatkan input yang ada dengan cara yang relatif cepat). Metode *naïve bayes* memanfaatkan teori probabilitas yang dikemukakan oleh ilmuan inggris *Thomas Bayes*, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. Tujuan algoritma NBC (Kusumadewi, 2009) adalah untuk melakukan klasifikasi pada data di kelas tertentu. Kelebihan *Naïve Bayes Classsifier* adalah sederhana tetapi memiliki akurasi yang tinggi (Rish, I., 2006, yang dikutip Rosdiansyah (2012)).

Dalam metode IR konsep seperti ini biasa di tandai dengan adanya satu set data yang di bagi menjadi dua bagian, yaitu data training dan data testing (Rozaq, Arifin, & Purwitasari, 2014). Sekelompok data yang akan diproses dan dicari kelasnya disebut dengan data testing, sedangkan data training merupakan data yang telah dihitung sebelumnya yang kemudian dibandingkan nilainya dengan sejumlah fitur yang ada di dalam data testing.

Naive Bayes Classifier menggunakan pendekatan *Teorema Bayes* untuk menghitung probabilitas kategori berdasarkan dokumen yang telah diketahui (Gilang Jalu Selo, 2013). Perhitungan nilai probabilitas tersebut menggunakan persamaan:

UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
© Hak Cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

$$P(V_j) = \frac{|docs\ j|}{|contoh|} \quad (2.1)$$

Keterangan:

$P(V_j)$: Peluang kemunculan dokumen yang memiliki kategori j.

$|docs_j|$: Jumlah dokumen setiap kategori j.

$|contoh|$: Jumlah dokumen dari semua kategori.

Dan terakhir melakukan perhitungan probabilitas setiap kata pada data uji terhadap data uji pada setiap kelas j, dengan cara:

$$\mathrm{P}(X_i|V_j) = \frac{n_k+1}{\mathrm{n}+|\mathrm{kosakata}|} \quad (2.2.)$$

Keterangan:

$\mathrm{P}(X_i|V_j)$: Peluang kemunculan X_i pada kategori V_j.

n_k : Jumlah frekuensi kemunculan setiap kata.

n : Jumlah frekuensi kemunculan kata dari setiap kategori.

|kosakata| : Jumlah semua kata dari semua kategori.

2.6 *White Box*

White box atau *Glass Box testing* adalah pengujian yang didasarkan pada pengecekan terhadap detil perancangan, menggunakan struktur kontrol dari desain program secara prosedural untuk membagi pengujian ke dalam beberapa *Test Case* (Nidrha & Dondeti, 2012). Penentuan kasus uji disesuaikan dengan struktur aplikasi, pengetahuan mengenai program digunakan untuk mengidentifikasi kasus uji tambahan.

Tujuan penggunaan *white box* untuk menguji semua statement program. Penggunaan metode pengujian *white box* dilakukan untuk :

1. Memberikan jaminan bahwa semua jalur independen suatu modul digunakan minimal satu kali.
2. Menggunakan semua keputusan logis untuk semua kondisi *true* atau *false*.

© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

3. Mengeksekusi semua perulangan pada batasan nilai dan operasional pada setiap kondisi.
4. Menggunakan struktur data internal untuk menjamin validitas jalur keputusan.

2.7 *Confusion Matrix*

Confusion Matrix merupakan metode untuk menghitung tingkat akurasi, dengan menghitung jumlah prediksi benar dan salah dari sebuah metode klasifikasi berbanding dengan data sesungguhnya atau prediksi target (Junaidi, 2015).

Menggunakan *Matrix* (NxN) dimana N adalah jumlah kelas. Berikut ini adalah contoh perhitungan akurasi dengan *Confusion Matrix* menggunakan dua kelas yaitu positif dan negatif.

Tabel 2. 1 Contoh *Confusion Matrix*

Confusion Matrix		**Target**			
		Positif	Negatif		
Model	Positif	A	b	*Positive Predictive Value*	a/(a+b)
	Negatif	C	D	*NegativePredictive Value*	d/(c+d)
		Sensitivity	*Specificity*	**Accuracy** = (a+d)/(a+b+c+d)	
		a/(a+c)	d/(b+d)		

Keterangan :

Accuracy : Jumlah prediksi benar (a+d) / Jumlah seluruh data (a+b+c+d)

Positive Predictive Value / Precision **:** Jumlah prediksi benar positif (a) / Jumlah target data positif (a+b)

Negative Predictive Value **:** Jumlah prediksi benar negatif (d) / Jumlah target data negatif (b+d)

Sensitivity /Recall **:** Jumlah prediksi benar positif (a) / Jumlah seluruh prediksi benar (a+c)

Specificity **:** Jumlah prediksi salah negatif (d) / Jumlah seluruh prediksi salah (c+d)

UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

2.8 Penelitian Terkait

Pada tabel 2.1 berikut ini menjelaskan tentang penelitian terkait yang telah dilakukan sebelumnya.

Tabel 2. 2 Penelitian Terkait

No	Nama penulis	Judul	Thn	Metode	Hasil
1	Ferly Gunawan, Mochammad Ali Fauzi, Putra Pandu Adikara	Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan *Naïve Bayes* dan Normalisasi Kata Berbasis *Levenshtein Distance*	2017	*Naïve bayes classifier*.	Dari hasil penelitian didapatkan tingkat akurasi tertinggi 96.9%.
2	Lutfi Budi Ilmawan, Edi Winarko	Aplikasi Mobile untuk Analisis Sentimen pada *Google Play*	2016	N*aïve bayes classifier*	Nilai akurasi dari *Naïve Bayes classifier* dari aplikasi yang dibangun sebesar 83,87%
3	Surya Hanggara, Tb Mohammad Akhriza, Mochammad Husni	Aplikasi Web Untuk Analisis Sentimen Pada Opini Produk Dengan Metode *Naïve Bayes Classifier*	2017	N*aïve bayes classifier*	*Naïve bayes* dapat dijadikan metode klasifikasi untuk analisis sentimen dengan tingkat akurasi rata-rata 70,43%-88,03%.
4	Azam Faiz Kamal, Budi Widjajanto	*Text Mining* Untuk Analisa Sentimen Ekspedisi Jasa Pengiriman Barang Menggunakan Metode *Naïve Bayes* Pada Aplikasi J & T Express.	2016	*Naïve bayes classifier*	Metode *Naïve Bayes* layak digunakan untuk pengklasifikasian sentiment ekspedisi jasa pengiriman barang terhadap 300 komentar dengan tingkat akurasi sebesar 83%
5	Putri Rizqiyah	Analisa Sentimen Menggunakan *Naïve Bayes* Untuk Mengetahui Presentase Komentar Pada Aplikasi GO-JEK	2016	*Naïve bayes classifier*.	*Sentiment Analysis* dapat dilakukan untuk mereview sebuah Aplikasi GO-JEK dengan memanfaatkan komentar pada aplikasi GO-JEK
6	Faishal Nuruz Zuhri, Andry Alamsyah	Analisis Sentimen Masyarakat Terhadap Brand SmartFren	2017	N*aïve bayes classifier*	Analisis Sentimen terhadap data kaskus mengenai8 brand Smartfren dilakukan dengan metode *Naïve*

No	Nama penulis	Judul	Thn	Metode	Hasil
		Menggunakan *Naïve Bayes Classifier* Di Forum kaskus			*Bayes Classifier* dengan menggunakan 1000 data latih didapatkan precision sebesar 98.42% recall sebesar 98.40% dan tingkat akurasi sebesar 98.40%. hal ini menunjukkan bahwa dengan metode *Naïve Bayes* dapat memisahkan data bersentimen positif dan negatif dengan sangat baik.
7	Mochammad Zoqi Sarwani dan Wayan Firdaus Mahmudy	Analisis Twitter Untuk Mengetahui Karakter Seseorang Menggunakan Algoritma *Naïve Bayes Classifier*	2015	*Naïve bayes classifier.*	Proses Klasifikasi menggunakan algoritma *Naïve Bayes Classifier* dengan 10 pengguna twitter sebagai data latih dan 10 pengguna twitter data uji, hasil yang diperoleh dibandingkan dengan data dari pakar sehingga didapatkan keakurasian data mencapai 100%.Proses Klasifikasi menggunakan algoritma *Naïve Bayes Classifier* dengan 10 pengguna twitter sebagai data latih dan 10 pengguna twitter data uji, hasil yang diperoleh dibandingkan dengan data dari pakar sehingga didapatkan keakurasian data mencapai 100%.
8	Arif Kurniawan	Text Mining Menggunakan Metode *Naïve Bayes Classifier* Untuk Analisa Sentimen Jasa Ekspedisi Barang	2017	*Naïve Bayes Classifie*	Analisa sentimen dengan menggunakan algoritma *Naïve Bayes Classifier* didapatkan akurasi sebesar 90.90% .
9	Zia Ayu Nuansa Gumilang	Implementasi *Naïve Bayes Classifier* dan Asosiasi Untuk Analisis Sentimen Data Ulasan Aplikasi E-Commerce Shopee Pada Situs *Google Play*	2018	*Naïve Bayes Classifier* dan *Asosiasi*	Hasil klasifikasi sentimen menggunakan model *Naïve Bayes Classifier* didapatkan tingkat akurasi sebesar 97,4%.

© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Nama penulis	Judul	Thn	Metode	Hasil
10	Ro'i Fahreza Nur Firmansyah,M.Ali Fauzi,Tri Afirianto	Sentiment Analisis Pada Review Aplikasi Mobile Menggunakan Metode *Naïve Bayes* dan *Query Expansion*	2016	*Naïve Bayes Classifier dan Query Expansion*	Berdasarkan hasil pengujian yang telah dilakukan penggunaan *Naïve Bayes-Query Expansion* dalam *Sentiment Analysis* terbukti lebih baik dengan akurasi sebesar 98% dibandingkan dengan *Naïve Bayes* tanpa menggunakan *Query Expansion* yang hanya mempunyai akurasi sebesar 95%.
11	Laila Ratnasari Putri, Mohamad Syahrul Mubarok,Adiwijaya	Klasifikasi Sentimen Ulasan Buku Berbahasa Inggris Menggunakan Information Gain dan *Naïve Bayes*	2017	*Naïve Bayes Classifier dan Information Gain*	klasifikasi sentimen pada ulasan buku berbahasa inggris menggunakan *Information Gain* dan *Naive Bayes* dari rata-rata *F1-score* menggunakan *5-fold-cross validation* mempunyai akurasi sebesar 88,28%
12	Fauziah Afshoh	Analisa Sentimen Menggunakan *Naïve Bayes* Untuk Melihat Persepsi Masyarakat Terhadap Kenaikan Harga Jual Rokok pada Media Sosial Twitter	2017	*Naïve Bayes Classifier* dan *Lexicon Based*	Menganalisa persepsi masyarakat terhadap kenaikan harga rokok pada media sosial *twitter*, *naïve Bayes Classifier*, hasil dari penelitian menunjukan bahwa dengan menggunakan data opini sebanyak 350 buah diketahui bahwa nilai sentimen positif yang paling banyak terbentuk dalam menanggapi topik kenaikan harga jual rokok dengan persentase *tweets* sebesar 53%

© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU
Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.