

## BAB II

### TINJUAN PUSTAKA DAN LANDASAN TEORI

#### 2.1 Tinjauan Pustaka

*Twitter* merupakan salah satu media sosial terpopuler saat ini yang banyak digunakan oleh semua kalangan usia kerana fungsi *twitter* yang sangat beragam contohnya sebagai tempat menuangkan perasaan, tempat membaca berita ataupun sebagai alat penyebar informasi. Seseorang bukannya hanya memakai media sosial ini sebagai media untuk berkomunikasi atau menuangkan perasaan, atau berpendapat tentang suatu hal-hal atau kejadian yang sedang dialami oleh seseorang tersebut. banyak kalangan anak muda yang suka membagikan perasaan mereka di *twitter* tentang keseharian baik di rumah atau pun di kampus. Kalangan mahasiswa biasanya sangat banyak membagikan moment-moment yang mereka lewati di kampus dengan memasukan foto mereka ke *Twitter* dan memberi keterangan di foto tersebut atau yang bisa kita sebut dengan *tweet*. *Tweet* tersebut akan menjelaskan keadaan apa yang dialami oleh orang tersebut, dari *tweet* itu akan menghasilkan suatu sentimen yang akan dianalisa apakah *tweet* tersebut bersifat, positif, netral, atau negatif. Setiap mahasiswa yang berkuliah pastinya tidak memiliki perasaan yang sama tiap harinya. Dari *tweet* seseorang inilah dilihat bisanya seseorang menulis opininya itu bersifat negatif atau netral.

Penelitian sebelumnya yang berisi Pemetaan Gaya belajar Mahasiswa dan Kecendruangan Emosi pada *Twitter*. Penelitan yang dilakukan penulis berkaitan dengan sentimen analisis pada *twitter*. Apabila seseorang menyampaikan opini atau gagasan mereka di *twitter* bisa dipastikan *tweet* orang tersebut mempunyai sentimen baik itu negatif, positif, ataupun netral. Penulis yang meneliti kasus ini bernama Robert Habibi memakai gaya belajar *visual, aural, read/write, kinastetic (VARK)*. Penulis Robert Habibi memakai

metode *Backpropagation*. Metode ini memakai bahasa pemrograman *java*, *Mysql* sebagai Database nya dan memakai *Weka* sebagai aplikasinya [5].

Pada penelitian yang berjudul analisis sentimen tentang kasus layanan akademik di perguruan tinggi, menjelaskan kasus keluhan yang sering terjadi di perguruan tinggi adalah kualitas layanan yang sering dikritisi oleh mahasiswa. Kualitas layanan ini sangat erat dengan kepuasan konsumen , dan menjadi masalah yang harus dipecahkan. Dipenelitian ini menjelaskan klasifikasi yang digunakan adalah *naïve bayes*, sedangkan untuk menguji dalam pengambilan keputusan yang akurat menggunakan metode deskriptif dan kemudian untuk teknis analisis data menggunakan metode *Analysis cross Validation* [6].

Penelitian yang berjudul Pengembangan Sosial Intelegensi Bisnis Akademik memanfaatkan media sosial *Twitter*. Intelegensi bisnis Akademik merupakan suatu sistem yang harus dikembangkan karena dapat membantu meningkatkan proses bisnis akademik di Universitas. Untuk membangun sistem ini biasanya harus memerlukan sejumlah data yang nantinya akan dianalisis terlebih dahulu. Data tersebut berasal dari *internal* dan *eksternal*. Data *internal* berasal dari dalam universitas dan data *eksternal* adalah data yang dikumpulkan dari media sosial *Twitter*. Penulis penelitian ini memakai *SQL Server* dan *Microsoft Visual studio 2008* sebagai aplikasi tempat pengolahan data [7].

Pada penelitian yang berjudul Penggunaan Media Sosial untuk Mengukur Kepuasan Pelanggan dalam Layanan Transit Publik. Penulis mengumpulkan pendapat pelanggan dalam sistem angkutan umum adalah dengan mengelola survei kepuasan pelanggan (*CSat*). *Platform* media sosial online menyediakan saluran komunikasi dua arah antara pelanggan dan organisasi, menawarkan lebih banyak kemampuan dari pada *survei CSat* yang biasa digunakan. Misalnya, media sosial online menyediakan aksesibilitas gratis ke data dan membuatnya layak untuk mempelajari dinamika diskusi online. Studi ini menganalisis diskusi online di *Twitter* tentang layanan transportasi umum untuk mengembangkan alat yang melengkapi *survei CSat*

transit. Analisis ini memiliki tiga komponen analisis: semantik, jejaring sosial, dan sentimen. Hasilnya dibandingkan dengan yang untuk *survei CSat* yang diterbitkan oleh agen transit Kanada terpilih untuk menentukan apakah media sosial dapat digunakan untuk memperkirakan perubahan dalam kepuasan pelanggan [8].

Berikut adalah Tabel perbandingan analisis sentimet yang pernah ada. Dapat dilihat pada tabel 2.1.



**Tabel 2.1. Perbandingan Aplikasi Twitter Analisis**

Penelitian	Habibi (2016)	Elfrida Cristella (2018)	Handianto (2016)	Moein Hosseini (2016)
Content	Pemetaan Gaya Belajar Mahasiswa dan Kecendrungan Emosi pada <i>Twitter</i>	Analisis Sentimen Tentang Kasus Layanan Akademik di Perburuan Tinggi (Telkom University)	Pengembangan Sosial Intelegensi Bisnis Akademik memanfaatkan media sosial <i>Twitter</i>	Penggunaan Media Sosial untuk Mengukur Kepuasan Pelanggan dalam Layanan Transit Publik
Metode	<i>backpropagation</i>	<i>naive bayes</i>	literature	Sematic analisis, sosial <i>network</i> dan sentimen analisis, <i>Sentistrength</i>
Database	<i>mysql</i>	Ms.Acc	<i>SQL Server, apache hadoop</i>	Ms. Acc
Peralatan	<i>Weka Workbench</i>	Kuisisioner, <i>Weka Workbench</i>	<i>Microsoft Visual Studio 2008 Business Intelligence studio</i>	Kuisisioner, <i>JetBrains PyCharm Community Edition</i>
Platfrom	Dekstop	Dekstop	Dekstop	Dekstop

## 2.2 LANDASAN TEORI

### 2.2.1 Sentimen Analisis

Sentimen analisis adalah suatu metode yang sering digunakan orang untuk melakukan identifikasi terhadap suatu kalimat, yang berujuk kepada suatu topik yang sedang dibicarakan. Hasil dari sentimen analisis ini adalah berupa nilai yang bersifat positif, negatif dan netral [9].

### 2.2.2 Text Mining

*Text mining* adalah proses pengeksploasian data yang berupa data text yang tidak terstruktur. Penambangan ini dilakukan oleh sebuah *software* pada komputer yang meminta sebuah masukan berupa kata kunci, sehingga kata kunci ini akan menghasilkan data text baru yang terstruktur [10].

### 2.2.3 Preprocessing

Pada proses ini data yang berhasil ditambang akan diolah lagi atau lebih tepatnya data *text* tersebut dibersihkan untuk membuang informasi-informasi yang tidak dibutuhkan. Berikut adalah langkah Preprocessing [11]:

#### a. Stopwords

*Stopwords* adalah proses pembersihan data *text* dengan menghilangkan tanda-tanda baca pada text yang tidak memiliki arti misalnya koma (,), hastag (#), mention (@).

#### b. Case Folding

*Case folding* adalah proses dimana mengubah semua data *text* yang memiliki huruf besar akan diubah menjadi huruf kecil.

### 2.2.4 Service Quality UAJY

Universitas Atma Jaya Yokyakarta adalah salah satu Universitas swasta terbaik dikota Jogja. Universitas Atma Jaya Yokyakarta sudah mendapat akreditasi A dari BAN-PT [12]. Selain dari segi akademik yang berprestasi

Universitas Atma Jaya juga berprestasi dalam bidang non akademik, bisa dilihat dari cabang olahraga bola Basketnya yang terlihat beberapa kali menjuari Liga Bola Basket tingkat Provinsi [13].

Akan tetapi UAJY juga harus memperhatikan *Service Quality* yang dilakukan terhadap mahasiswanya untuk mendukung proses belajar dan mengajarnya. *Service Quality* ini dibagi kedalam 5 Dimensi pokok antara lain: Dimensi *Tangibles*, Dimensi *Reliability*, Dimensi *Responsiveness*, Dimensi *Assurance*, dan Dimensi *Empathy* [14].

**1. Dimensi Tangibles**

Berkenaan dengan daya tarik fasilitas fisik, peralatan/perlengkapan yang lengkap, dan material yang digunakan perusahaan bersih, serta penampilan dari karyawan rapi.

**2. Dimensi Reliability**

Dimensi (*Reliability*) Berkaitan dengan kemampuan perusahaan untuk memberikan pelayanan yang akurat sejak pertama kali tanpa melakukan kesalahan apapun dan menyampaikan jasanya sesuai dengan waktu yang disepakati.

**3. Dimensi Responsiveness**

Daya Tanggap (*Responsiveness*) Berhubungan dengan kesediaan dan kemampuan karyawan untuk membantu para konsumen dan merespon permintaan mereka, serta menginformasikan kapan jasa akan diberikan dan kemudian memberikan jasa secara cepat.

**4. Dimensi Assurance**

Jaminan (*Assurance*) perilaku karyawan yang mampu menumbuhkan kepercayaan konsumen terhadap perusahaan dan perusahaan bisa menciptakan rasa aman bagi para konsumennya. Jaminan juga berarti bahwa para karyawan selalu bersikap sopan dan menguasai pengetahuan dan keterampilan yang dibutuhkan untuk menangani setiap pertanyaan atau masalah konsumen.

## 5. *Dimensi Emphaty*

Menyatakan bahwa perusahaan memahami masalah para konsumennya dan bertindak demi kepentingan konsumen, serta memberikan perhatian personal kepada para konsumen dan memiliki jam operasi yang nyaman [15].

### 2.2.5 Twitter

*Twitter* adalah salah satu media sosial yang paling populer saat ini yang memungkinkan kita untuk dapat mengirimkan pesan, menerima pesanya yang memiliki *teks* maksimal 140 karakter. Fitur yang terdapat pada *Twitter* antara lain sebagai berikut:

a. Halaman Utama (*Home*)

Pengguna bisa melihat berbagai *tweet* dari orang-orang yang telah diikuti

b. Profil pengguna

Profil berfungsi sebagai menampilkan berbagai informasi-informasi yang tulis oleh pemilik akun tersebut, sehingga orang-orang dapat melihatnya.

c. Pengikut (*Followers*)

Pengikut adalah orang-orang telah mengikuti pengguna, sehingga *followers* dapat melihat setiap *tweet* yang ditulis oleh pengguna.

d. Mengikuti (*Following*)

*Following* adalah pengguna yang megikuti pemilik akun lain yang ada ditwitter, sehingga pengguna tersebut dapat melihat *tweet* orang-orang yang telah diikuti.

e. Pesan langsung (*Direct Messege*)

Memungkinkan pemilik akun *twitter* dapat mengirimkan pesan secara langsung kepada pengguna *Twitter* yang lainnya.

f. Pagar (#) / (*Hastag*)

Merupakan simbol seperti menyerupai pagar yang berfungsi memudahkan seseorang dalam melakukan pencarian suatu topik.

g. *Mentions* (@)

*Mentions* berfungsi sebagai apabila ada pengguna *Twitter* yang ingin melakukan percakapan dengan orang lain dihalaman utama, pengguna tersebut tinggal memakai tanda *mentions* ke orang yang dituju.

h. Trending Topik

Fitur ini berfungsi merangkum semua *tweet-tweet* yang sedang hangat diperbincangkan oleh pengguna lainnya. Apa bila terdapat banyak *tweet* yang menggunakan tanda (#) pada suatu kalimat tertentu, itu menandakan topik yang mereka bahas sedang menjadi trending topik.

### 2.2.6 Metode *Sentistrength*

Metode *SentiStrength* adalah suatu metode yang dapat mendapatkan skor kekuatan sentimen dan dapat mengelompokkan sentimen atas positif, netral, dan negatif sesuai dengan nilai-nilai yang terkandung dari *tweet* yang disampaikan [16]. Metode ini mengukur nilai sentimen dari kamus-kamus yang sudah diberi bobot. Bobot yang diberi nilai -5 (memiliki sentimen negatif yang sangat kuat) sampai 5 (memiliki nilai sentimen positif yang sangat kuat). Kamus-kamus ini didapat dari hasil translasi kamus sentimen berbahasa Inggris yang melibatkan 3 responden ahli, mahasiswa S2 Linguistik Universitas Gadjah Mada (UGM) untuk memberi nilai bobot pada sentimen.

a. Kamus Sentimen

Di dalam kamus ini terdapat beberapa kata yang sudah memiliki bobot atau nilai sentimennya. Berikut adalah contoh kata-kata yang sudah diberi bobot dapat dilihat pada tabel 2.2.

**Tabel 2.2. Contoh kata kamus sentimen**

<b>Kata</b>	<b>Bobot</b>
Aman	4
Anarki	-5
Baik	4
Celakalah	-5
Dicemburui	2
Dihargai	4

b. Kamus Idiom

Kamus idiom merupakan suatu kata ungkapan yang memiliki arti yang berbeda dari kata aslinya [15]. Berikut adalah contoh kata-kata idiom yang sudah diberi bobot dapat dilihat pada tabel 2.3.

**Tabel 2.3. Contoh kata idiom dalam kamus sentimen**

<b>Kata</b>	<b>Bobot</b>
Cinta mati	5
Jatuh hati	4
Kambing hitam	-4
Cari perhatian	-4
Buah hati	1

c. *Boosterwords*

*Boosterwords* adalah kata yang dapat meningkatkan atau mengurangi nilai makna sentimen dari suatu kata yang ada disebelahnya [15]. Berikut adalah contoh kata-kata yang sudah diberi bobot dalam kamus sentimen dapat dilihat pada tabel 2.4.

**Tabel. 2.4. Contoh kata Boosterwords dalam kamus sentimen.**

<b>Kata</b>	<b>Bobot</b>
Betul-betul	1
Paling	1
Sempurna	1
semakin	1
kurang	-2

d. Kata negasi

Kata negasi adalah kata yang dapat mengubah arti dari suatu kalimat contohnya kata "pantang" yang diketahui kata "pantang" adalah suatu kata negatif, akan tetapi jika didahului dengan kata "jangan pantang menyerah" sehingga bisa mengubah nilai sentimennya menjadi kalimat yang positif. Berikut adalah contoh kalimat kata negasi dalam kamus sentimen dapat dilihat pada tabel 2.5.

**Tabel 2.5. Contoh kata negasi dalam kamus sentimen.**

<b>No</b>	<b>Kata</b>
1	Jangan
2	Tak
3	Gak
4	tanpa
5	bukan

e. Kata Tanya

Kata Tanya ini dapat mengubah suatu makna dari suatu kalimat misalnya terdapat kata "Pencuri" pencuri dalam suatu kalimat, tetapi kalimat tersebut belum bisa dikatakan sebagai kalimat negatif karena ada beberapa kata pendukung seperti kata tanya

berikut “ Apakah orang itu seorang Pencuri?” dengan adanya kata Tanya bisa mengubah kalimat tersebut, bukan positif, bukan negatif, melainkan kalimat netral. Berikut adalah contoh kata Tanya dalam kamus sentimen dapat dilihat pada tabel 2.6.

**Tabel 2.6. Contoh kata Tanya dalam kamus sentimen.**

No	Kata
1	Kenapa
2	Mengapakah
3	Apakah
4	Salakah
5	benarkah

#### **2.2.6 Korpus**

Menurut buku *Corpus Linguistics* Korpus didefinisikan sebagai koleksi atau sekumpulan contoh teks tulis atau lisan dalam bentuk data yang dapat dibaca dengan menggunakan seperangkat mesin dan dapat diberi catatan berupa berbagai bentuk informasi linguistik. Korpus adalah sekumpulan teks-teks atau kata-kata yang membentuk suatu kamus [17].

Korpus dirangkum dari berbagai sumber seperti media cetak dan elektronik dirangkum yang didalamnya pasti mengandung kata frasa, klausa dan wacana sehingga menjadi sebuah kamus yang dapat digunakan [18]. Korpus juga sering disebut sebagai bank bahasa karena terdapat data yang sangat besar.

#### **2.2.7 N-gram**

N-gram adalah potongan n suatu karakter dalam string tertentu atau potong n kata dalam kata tertentu [19]. Berikut adalah contoh dari N-gram yang berasal dari kata “TUGAS” dapat dilihat pada tabel 2.7.

**Tabel 2.7. Contoh potongan N-gram dari kata TUGAS**

<b>Nama</b>	<b>N-gram Karakter</b>
Uni-gram	T, U, G, A, S
Bi-gram	_T, TU, UG, GA, AS, S_
Tri-gram	_TU, TUG, UGA, GAS, AS_, S_
Quad-gram	_TUG, TUGA, UGAS, GAS_,AS_,S_

Karakter “\_” berfungsi sebagai spasi yang berada diawal atau diakhir. Berikut adalah contoh N-gram pada sebuah kalimat “saya berhasil mengerjakan skripsi saya selama satu semester” dapat dilihat pada tabel 2.8

**Tabel 2.8. Contoh potongan N-gram pada suatu kalimat.**

<b>Nama</b>	<b>N-gram kata</b>
<i>Uni-gram</i>	Saya, berhasil, mengerjakan, skripsi, saya selama, satu, semester
<i>Bi-gram</i>	saya berhasil, berhasil mengerjakan, mengerjakan skripsi, skripsi saya, saya selama, selama satu, satu semester.
<i>Tri-gram</i>	saya berhasil mengerjakan, berhasil mengerjakan skripsi, mengerjakan skripsi saya, skripsi saya selama, saya selama satu, selama satu semester
<i>Quad-gram</i>	saya berhasil mengerjakan skripsi, saya selama satu semester