

## **BAB 3**

### **LANDASAN TEORI**

#### **3.1. Brand**

Berdasarkan *American Marketing Association*, *brand* didefinisikan sebagai nama, istilah, desain, simbol atau ciri lain yang dapat mengidentifikasi produk atau jasa penjual dibandingkan dengan penjual lain (*American Marketing Association*, 2015). *Brand* secara konseptual merupakan satu bagian keuntungan *tangible* dan *intangible* yang dapat meningkatkan daya tarik produk atau jasa, di atas nilai fungsionalitasnya (Maurya & Mishra, 2012). Dari definisi tersebut dapat dijelaskan bahwa *brand* merupakan identifikasi khusus yang membuat pelanggan dapat mengenalinya. Identifikasi khusus tersebut penting bagi penjual untuk meningkatkan kesadaran pelanggan atas produk atau jasa yang mereka jual. Kesadaran pelanggan dapat ditingkatkan dengan menerapkan kebijakan pemasaran atau *branding*. Hal tersebut akan mengenalkan produk atau jasa dari penjual ke pasar yang lebih luas.

*Branding* merupakan proses menciptakan nilai melalui penyediaan tawaran yang menarik dan konsisten sehingga memberikan pelanggan pengalaman yang memuaskan dan membuat mereka datang kembali. Dengan mengembangkan kepercayaan kepada *brand* melalui kepuasan penggunaan dan pengalaman kepada pelanggan membuat perusahaan memiliki kesempatan untuk membangun hubungan dengan pelanggan untuk memperkuat *brand* mereka (Simmons, 2007).

*Brand* juga merupakan representasi dari pengalaman pelanggan dengan produk atau jasa (*American Marketing*

Association, 2015). Dari pengalaman tersebut pelanggan memberikan opini positif atau negatif. Opini pelanggan terhadap *brand* berasal dari kebijakan yang diterapkan kepada produk atau jasa oleh pemilik *brand*. Hasil dari kebijakan tersebut dapat membuat pelanggan mengutarakan opini mereka terhadap *brand*. Opini tersebut akan mempengaruhi proses bisnis dari *brand*. Penting bagi penjual untuk bijaksana dalam mengeluarkan kebijakan. Karena sentimen positif pelanggan terhadap *brand* secara tidak langsung dapat meningkatkan pendapatan dan proses bisnis, sedangkan sentimen negatif secara tidak langsung dapat menurunkan pendapatan atau proses bisnis (Pullig, 2008).

### **3.2. Jejaring Sosial**

Situs jejaring sosial mulai diciptakan pada medio 1990-an. Beberapa contoh jejaring sosial yang muncul pada medio 1990-an adalah *Six Degrees*, *BlackPlanet*, *Asia Avenue*, dan *MoveOn*. Situs jejaring sosial online ini sudah dapat membuat individu berinteraksi. Pada medio 2000-an media sosial berkembang pesat dengan semakin banyak situs jejaring sosial bermunculan. Hal ini mengubah interaksi komunikasi individu dan organisasi konvensional menjadi berdasarkan jejaring sosial. Perubahan ini membuat individu dan organisasi dapat membangun model relasi baru di mana tidak ada batasan jarak dan waktu untuk melakukan interaksi komunikasi.

Jejaring sosial merupakan layanan situs berbasis *website* yang memungkinkan individu membangun profil publik atau semi publik yang terikat dalam sistem, membuat dan melihat catatan mengenai kepada siapa saja pengguna berbagi koneksi (Ellison & Boyd, 2007). Dari

penjelasan tersebut dapat disimpulkan bahwa jejaring sosial merupakan sarana komunikasi terkomputerisasi antar individu yang terhubung dengan jaringan internet. Dengan jejaring sosial hubungan komunikasi antar individu semakin berkembang pesat. Setiap individu dapat membangun relasi dengan berbagai individu dari berbagai belahan dunia melalui jejaring sosial dan penggunaan waktu yang lebih fleksibel.

### **3.3. Data Mining**

Data *mining* merupakan proses ekstraksi informasi dan pengetahuan yang secara potensial berguna dan orang-orang tidak mengetahuinya, serta ekstraksi dilakukan dari sekumpulan data yang besar, tidak lengkap, *noise*, dan acak (Sahu, et al., 2011). Data *mining* merupakan langkah penting dalam pencarian pengetahuan dalam basis data yang menghasilkan pola yang bermanfaat atau model dari data (Silwattananusarn & Tuamsuk, 2012). Dapat disimpulkan bahwa data *mining* merupakan proses pencarian pengetahuan berupa pola atau model data yang bermanfaat dari basis data yang besar dan kotor.

Data *mining* merupakan penggabungan dari 3 teknologi yaitu kekuatan komputasi, basis data dan algoritma pembelajaran dan statistik (Sahu, et al., 2011). Kekuatan komputasi diperlukan untuk mempercepat proses data *mining* karena terdiri dari jumlah data yang besar. Pengetahuan basis data diperlukan untuk mengelola data dalam bentuk yang tepat sehingga mempermudah proses data *mining*. Untuk mendapatkan pola pengetahuan yang berguna, digunakan algoritma pembelajaran dan statistik. Peningkatan dan perbaikan dari ketiga teknologi ini akan menghasilkan hasil yang

lebih akurat dan waktu proses yang lebih efisien.

Secara umum, data *mining* memiliki 5 tahapan pengerjaan, yaitu seleksi data, data *preprocessing*, transformasi data, data *mining* dan interpretasi atau evaluasi. Seleksi data merupakan pemilihan sumber data yang relevan untuk dianalisa. Data *preprocessing* merupakan tahapan pembersihan data dari *noise* dan tidak konsisten. Transformasi data adalah tahapan untuk membentuk sumber data yang digunakan menjadi bentuk yang tepat untuk data *mining*. Data *mining* adalah proses untuk mendapatkan pola informasi dan pengetahuan dari sumber data menggunakan algoritma data *mining*. Interpretasi atau evaluasi merupakan tahapan untuk menganalisa pola yang di dapat dalam bentuk pengetahuan untuk direpresentasikan dalam bentuk yang mudah dipahami.

Dalam tahap *preprocessing* berbentuk teks pembersihan data dilakukan dengan mengenali setiap kata yang dimasukkan. Pengenalan kata dilakukan dengan proses *Stemming*. *Stemming* adalah proses pemetaan dan penguraian berbagai bentuk (variasi) dari suatu kata menjadi bentuk kata dasarnya (Indriyono, et al., 2015). Dalam kasus teks berbahasa Indonesia terdapat algoritma yang dapat digunakan adalah algoritma Nazief & Adriani (Nazief & Adriani, 1996). Algoritma ini akan mengubah setiap kata yang ditemukan menjadi kata dasar dan mengecek ketersediaannya dalam Kamus Besar Bahasa Indonesia (KBBI). Proses penemuan kata dasar dilakukan dengan menghilangkan akhiran dan awalan berdasarkan *Inflection Suffixes*, *Possesive Pronouns*, *Derivation Suffixes*, *Derivation Prefix*, dan aturan khusus kombinasi awalan-akhiran.

Salah satu algoritma klasifikasi data mining adalah *Naïve Bayes Classifier*. Metode ini merupakan metode kelompok Bayesian yang paling sederhana. Metode *Naïve Bayes Classifier* mengelompokkan data ke dalam kelas tertentu berdasarkan dari data *training* yang sudah jelas kelompok kelas datanya. Data *training* tersebut akan digunakan untuk menghitung jumlah kemiripan data yang dicari dalam kelas-kelas yang ada di data *training*.

$$P(C_i|\mathbf{X}) = \frac{P(\mathbf{X}|C_i)P(C_i)}{P(\mathbf{X})} \dots\dots\dots (1)$$

Persamaan model (1) memiliki arti yaitu  $P(C_i|\mathbf{X})$  merupakan probabilitas  $C_i$  terjadi jika  $\mathbf{X}$  sudah terjadi,  $P(C_i)$  adalah probabilitas  $C_i$  dalam data dengan sifat *independent* terhadap  $\mathbf{X}$ ,  $\mathbf{X}$  adalah kumpulan atribut,  $P(\mathbf{X}|C_i)$  adalah probabilitas  $\mathbf{X}$  terjadi jika  $C_i$  sudah terjadi berdasarkan data *training*. Dengan  $P(\mathbf{X})$  konstan maka dapat dilakukan penyederhanaan model menjadi :

$$P(C_i|\mathbf{X}) = P(\mathbf{X}|C_i)P(C_i) \dots\dots\dots (2)$$

Model (2) untuk mencari posisi probabilitas nilai dari data yang dicari berdasarkan data *training*. Secara umum model tersebut menunjukkan bahwa posisi probabilitas didapat dari nilai kemiripan data dibagi nilai data yang dicari (untuk nilai  $P(\mathbf{X}|C_i)$ ) dikalikan dengan nilai bukti dari data *training* (untuk nilai  $P(C_i)$ ). Untuk metode klasifikasi data *mining* sering terdapat banyak kelas yang digunakan untuk mendapatkan kesimpulan menuju kelas tujuan.

$$P(\mathbf{X}|C_i) = \prod_{k=1}^n P(x_k | C_i) = P(x_1 | C_i) \times P(x_2 | C_i) \times \dots \times P(x_n | C_i) \cdot \cdot \quad (3)$$

Model (3) untuk mendapatkan nilai probabilitas X terjadi jika  $C_i$  sudah terjadi adalah dapat dilakukan dengan mengalikan nilai posisi probabilitas dari masing-masing kelas yang mendukung kelas tujuan. Dari nilai-nilai itu akan dihitung nilai posisi probabilitas data yang dicari masuk ke kelas tujuan.

Nilai yang didapatkan dari model (3) akan digunakan untuk mendapatkan nilai akhir  $C_i$  terjadi jika X sudah terjadi pada model (2). Nilai posisi probabilitas  $C_i$  terjadi jika X sudah terjadi pada model (2) yang terbesar adalah hasil yang diambil untuk menunjukkan kelasnya.