

# BAB I PENDAHULUAN

## 1.1.Latar Belakang

Salah satu makanan pokok masyarakat Indonesia adalah jenis makanan yang diolah dari Unggas. Ayam merupakan makanan yang umum dijumpai sebagai menu makan keluarga hampir disemua kalangan ekonomi atas maupun bawah. Bahkan menu unggas ini juga digemari oleh semua umur karena rasanya yang enak serta nutrisi yang baik pada olahan daging unggas ini. Dilihat dari konsumsi masyarakat secara umum, daging unggas khususnya ayam merupakan jenis bahan olahan yang dikonsumsi sebagai menu harian di meja makan keluarga maupun di rumah makan pada umumnya.

Hal ini dapat dipelajari dari testimoni dan opini atau hasil review dari konsumen yang pernah mempunyai pengalaman pada Geprekan Jember tersebut sebelumnya. Beberapa review tentang Geprekan Jember dapat membantu konsumen dalam mengetahui kualitas Geprekan Jember. Saat ini konsumen yang menulis opini dan pengalaman secara online semakin meningkat. Membaca review tersebut secara keseluruhan dapat memakan waktu yang sangat lama, namun jika hanya sedikit review yang dibaca dievaluasi akan menjadi bias. Klasifikasi *sentiment* bertujuan untuk mengatasi masalah ini dengan secara otomatis mengelompokkan *review* pengguna menjadi opini positif, atau negatif (Zhang, Ye & Li, 2011). Untuk itu perlunya pengkajian ulang tentang review Geprekan Jember dengan cara klasifikasi review kedalam class positif, netral dan negatif agar pada akhirnya pemilik usaha dapat mengetahui tanggapan konsumen untuk melakukan perubahan dan memberikan kepuasan terhadap konsumen.

Untuk mempermudah pengolahan klasifikasi testimoni di Geprekan Jember penelitian memanfaatkan media SMS Gateway agar sistem dapat terintegrasi secara langsung. Salah satu software yang bisa digunakan untuk merealisasikan SMS Gateway adalah Gammu. Gammu merupakan aplikasi yang berbasis *opensource* yang digunakan sebagai konektor antara komputer dengan handphone. Kelebihan gammu yaitu dapat dijalankan di sistem operasi *Windows* maupun *Linux*,

banyak handphone yang support gammu, menggunakan *database MySql* dan bisa menggunakan *interface web-based*.

Analisa *sentiment* atau *opinion mining* adalah studi komputasi mengenai pendapat, perilaku dan emosi seseorang terhadap entitas. Entitas tersebut dapat menggambarkan individu, kejadian atau topik, tersebut kemungkinan besar dapat berupa review (Liu, Wang, etc, 2011). Teknik klasifikasi yang biasa digunakan untuk analisis sentiment review diantaranya *Naïve bayes*, *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (KNN) (Dehkharghani, Mercan, Javeed & Saygin, 2014).

Sentiment analysis sebagai bagian dari *text mining* mampu menjadi solusi dalam pengkategorian opini secara otomatis berdasarkan learning terhadap data yang telah dipersiapkan sebelumnya. Namun permasalahan besar yang muncul pada kasus text mining adalah tingginya dimensionalitas (*feature*). Dimensi yang tinggi, terlebih lagi dengan banyaknya yang tidak berkaitan dengan kategorisasi dapat menjadi noise pada data dan menurunkan tingkat akurasi dalam klasifikasi (Kiribiya, Frank, Pfahringer and Holmes, 2005.). *Multinomial Naïve Bayes* sebagai salah satu algoritma yang sering dipakai untuk kasus *text mining* menggunakan TF-IDF untuk melakukan *feature weighting*, namun ini masih dianggap kurang maksimal karena pada kasus text mining dikarenakan sifat aslinya yang digunakan pada data bertipe kategorikal. Pada sentiment analysis, setiap features memiliki pengaruh yang berbeda-beda terhadap hasil kategorisasi (Tan, 1999), karenanya diperlukan perubahan terhadap fitur ini agar mampu menyesuaikan terhadap kasus. Sentiment Analysis. Melihat hal tersebut, Tao Dong dalam papernya mengajukan sebuah algoritma baru turunan dari Gini Index yang dapat digunakan sebagai pengganti daripada algoritma IDF dalam kasus klasifikasi teks.

Algoritma *Cosine Similarity* digunakan untuk menghitung tingkat kesamaan antar data testing dengan data training testimoni yang memiliki *term* sama. Dari kelompok mentions tersebut, administrator akan memberikan solusi yang relevan terhadap keluhan. Selanjutnya ketika ada mentions baru berupa keluhan, sistem akan menampilkan rekomendasi solusi untuk kemudian dimoderasi oleh administrator. Dengan cara ini administrator tidak perlu memberikan jawaban yang

sama secara berulang-ulang kepada akun yang berbeda. Hal ini akan mempercepat kinerja administrator dalam memberikan tanggapan terhadap keluhan konsumen.

### **1.2.Rumusan Masalah**

Rumusan masalah dari uraian diatas adalah :

1. Bagaimana melakukan analisa terhadap *feedback* konsumen Geprekan Jember dengan menggunakan *tf-idf* dan *cosine similarity*?
2. Berapa tingkat akurasi metode *tf-idf* dan *cosine similarity* dalam klasifikasi *feedback* pelanggan?

### **1.3.Batasan Penelitian**

Agar penelitian ini tidak menyimpang dari rumusan masalah yang telah dijabarkan di atas maka penulis memberikan batasan penelitian sebagai berikut:

1. Data yang digunakan adalah *feedback*/testimoni pelanggan di Geprekan Jember.
2. Kata-kata (*word*) yang di proses adalah kata baku menurut KBBI.

### **1.4.Tujuan Penelitian**

Tujuan dari penelitian ini yaitu:

1. Mengukur tingkat rata-rata kecepatan sistem dalam klasifikasi *feedback* pelanggan terhadap Geprekan Jember menggunakan metode *tf-idf* dan *cosine similarity*.
2. Menghitung nilai tingkat akurasi metode *tf-idf* dan *cosine similarity* dalam mengklasifikasi *feedback* pelanggan di Geprekan Jember.

### **1.5.Manfaat Penelitian**

Manfaat yang diharapkan penulis adalah sebagai berikut:

1. Dapat membantu pemilik Geprekan Jember dalam memberikan pelayanan yang lebih dari hasil klasifikasi *feedback*/testimoni pelanggan.
2. Mengetahui klasifikasi *feedback* pelanggan dengan cepat.

3. Penulis dapat memahami model klasifikasi komentar pelanggan dengan menggunakan metode *tf-idf* dan *cosine similarity*.