



### **Hak cipta dan penggunaan kembali:**

Lisensi ini mengizinkan setiap orang untuk mengubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

### **Copyright and reuse:**

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Perkembangan aplikasi dalam mobile sangat meningkat. Pada awal 2019, dari survey yang dilakukan hootsuite We are social jumlah pengguna aktif internet di Indonesia sebanyak 96% (Hootsuite,2018). Survey lainnya dilakukan oleh App Annie ada 5.087 triliun yang melakukan pengunduhan aplikasi *mobile* (App Annie,2018). Dengan banyaknya pengguna *mobile*, maka diperlukannya juga aplikasi *mobile* yang memadai dengan kebutuhan *user* masing-masing. Setelah user mendapatkan aplikasi yang dibutuhkan, *user* akan mencoba aplikasi tersebut apakah sudah sesuai dengan kebutuhannya atau tidak.Saat *user* menggunakan aplikasi *mobile*, *user* akan mengalami pengalaman dalam menggunakannya dari fungsionalitas, *user interface* dan hal lainnya. *User* juga dapat memberi komentar atau saran pada aplikasi yang disebut dengan *user feedback*. *User feedback* yang masuk akan disampaikan pada *engineer* untuk evaluasi aplikasi dan membuat pengembangan aplikasi menjadi lebih baik.

Mengumpulkan kebutuhan atau keluhan *user* merupakan fokus *requirement engineering* (RE). *Requirement engineering* sebagian besar fokus untuk melibatkan pengguna sistem, mengumpulkan kebutuhan pengguna, dan mendapatkan *feedback* dari pengguna (Maalej et al. 2016). *User feedback* mendapatkan informasi seperti *bug report* yang menjelaskan kelemahan dari aplikasi yang digunakan dan juga bisa masukan dari pengguna dari apa yang kurang dari aplikasi yang digunakan (Maalej et al. 2016). Dengan adanya *user feedback*, *engineer* dari aplikasi tersebut dapat memahami kebutuhan dan keluhan dari pengguna. Setelah *engineer* mendapatkan *user*

*feedback*, maka akan dilakukan identifikasi dan pengklasifikasian *user feedback*. Untuk melakukan klasifikasi pada *user feedback* dapat menghabiskan waktu yang banyak, dikarenakan *user feedback* tidak sedikit, namun bisa ribuan *user feedback* setiap saat. Oleh karena itu, diperlukan adanya sistem otomatis untuk analisa sentimen *user feedback*. Untuk menganalisa opini dari *user feedback*, dapat menggunakan teknik analisis sentimen.

*Sentiment analysis (SA)* merupakan metode proses dalam memahami, meng-ekstrak, dan mengolah sebuah data yang berupa tekstual yang didapatkan dari kalimat opini yang bekerja otomatis sehingga didapatkannya informasi sentimen yang terkandung didalamnya (Somantri dan Apriliani, 2018). Analisis sentimen dilakukan sebagai upaya untuk melihat sebuah pendapat atau opini terhadap objek atau permasalahan oleh seseorang yang akhirnya nanti mempunyai nilai kecenderungan apakah menilai positif atau negatif (Somantri dan Apriliani, 2018). Namun, ada beberapa batasan yang membuat analisis sentiment dan *engineer* menggunakan kalimat opini dari *user feedback*. Karena, dari kalimat opini tersebut dapat memberikan saran atau masukan pada *engineer* untuk membuat aplikasi lebih baik. Dari kalimat opini yang sangat banyak dari berbagai *user* membuat upaya yang dilakukan pada analisis sentiment sangat besar. Lalu, kualitas dari *user feedback* atau kalimat opini yang diberikan *user* sangat bervariasi, dari saran yang membantu sampai ide yang inovatif. Selain penggunaan *user feedback* dalam evaluasi aplikasi rating pada aplikasi juga mewakili dari kalimat opini yang ada. Namun, kegunaan dari *rating* dalam *user feedback* untuk *engineer* sangat terbatas, karena *rating* mewakili rata-rata untuk keseluruhan aplikasi dan dapat menggabungkan keduanya antara evaluasi positif atau negatif dari satu fitur (Guzman dan Maalej, 2014). Dari hal yang sudah dibahas untuk mengklasifikasikan sentimen analisis dibagi menjadi 3 kategori yaitu positif, negative,

dan netral. Adapun penelitian terkait yang dilakukan dalam pengklasifikasian sentimen analisis.

Dalam penelitian yang terkait ada beberapa metode *machine learning* yang digunakan dalam analisis sentimen. (Somantri dan Apriliani, 2018) Memperbandingkan penerapan seleksi fitur untuk optimalisasi tingkat akurasi sentiment analisis klasifikasi rekomendasi menggunakan metode *Support Vector Machine* (SVM) dan didapatkan tingkat akurasi 72,45% dengan fitur seleksi *information gain*. (Chandani dan Purwanto, 2015) melakukan komparasi algoritma klasifikasi machine learning pada sentiment review film, SVM mendapatkan hasil terbaik dengan akurasi 81,10% dan penelitian tersebut mengatakan bahwa SVM merupakan algoritma yang paling baik.

Menurut Alfian Nur Rahmad dan Feddy Setio Pribadi (2015), Bahwa penggunaan fitur seleksi seperti *chi square* dalam klasifikasi dapat meningkatkan hasil baik recall, precision, F-measure, dan akurasi walaupun peningkatannya tidak terlalu signifikan, tapi dapat mempengaruhi hasil klasifikasi (Rahmad dan Pribadi, 2015). Penelitian tersebut mengatakan penggunaan *chi square* dalam pemilihan fitur dalam pengklasifikasikan berita menghasilkan akurasi klasifikasi yang lebih baik (Rahmad dan Pribadi, 2015). Penelitian Cahyono dan Saprudin melakukan analisis sentiment analisis *tweets* berbahasa sunda menggunakan *naïve bayes classifier* dengan *feature*

*selection chi square* mendapatkan akurasi sebesar 78,48 % (Cahyono and Saprudin 2019).

Berdasarkan penjelasan latar belakang yang sudah dijelaskan, penelitian yang akan dilakukan adalah menggunakan algoritma *support vector machine* untuk mengklasifikasikan *user feedback* dengan tambahan fitur seleksi *chi square*.

### **1.2. Rumusan Masalah**

Berdasarkan uraian pada latar belakang, maka rumusan masalah pada penelitian ini sebagai berikut.

1. Bagaimana mengimplementasikan metode Support Vector Machine dan Chi Square untuk menganalisa sentiment *user feedback*.
2. Bagaimana mengukur performa metode Support Vector Machine dan Chi Square yang digunakan untuk analisis sentimen *user feedback*.

### **1.3. Batasan Masalah**

Batasan masalah dalam penelitian ini sebagai berikut.

1. Sistem analisis sentimen ditunjukkan untuk *user feedback* berbahasa Indonesia.
2. Dataset yang digunakan diambil dari penelitian dengan judul “*An Experimental Study of Lexicon-based Sentiment Analysis on Bahasa Indonesia*” yang dilakukan oleh Endang Wahyu Pamungkas dan Divi Galih Prasetyo putri yang memiliki tiga kelas sentimen yaitu positif, negatif dan netral. Pelabelan pada dataset ini dilakukan menggunakan *SentiWordNet*.
3. Kernel yang digunakan pada algoritma *Support Vector Machine* adalah kernel Gaussian RBF.

#### **1.4. Tujuan Penelitian**

Tujuan dari penelitian ini sebagai berikut.

1. Mengimplementasikan algoritma Support Vector Machine dan Chi Square untuk analisis sentimen *user feedback*.
2. Mengevaluasi dalam mengukur performa dari metode analisis sentimen yang digunakan dengan cara melakukan pengujian terhadap *user feedback*.

#### **1.5. Manfaat Penelitian**

Manfaat dari penelitian ini adalah membantu untuk pengklasifikasian *user feedback* pada aplikasi, mempercepat klasifikasi *user feedback* secara otomatis. Dengan penelitian ini diharapkan dapat membuka kesempatan *developer* untuk lebih lanjut meneliti metode analisis sentiment *user feedback* berbahasa Indonesia.

#### **1.6. Sistematika Penulisan**

Sistematika penulisan laporan skripsi ini dapat dijabarkan dalam detail sebagai berikut:

##### **BAB I            LATAR BELAKANG**

Bab I Pendahuluan berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

##### **BAB II           LANDASAN TEORI**

Bab II Landasan Teori berisi penjelasan atau landasan teori mengenai *Sentimen Analisis, User Feedback, Data preprocessing, Support Vector Machine, Chi Square, Evaluasi Performa* sebagai algoritma yang digunakan untuk penelitian ini.

**BAB III      METODOLOGI PENELITIAN DAN PERANCANGAN SISTEM**

Bab III Metodologi Penelitian dan Perancangan Sistem berisi penjelasan mengenai metodologi penelitian yang digunakan. Bab ini juga berisi perancangan *flowchart*.

**BAB IV      IMPLEMENTASI DAN ANALISIS**

Bab IV Implementasi dan Analisis berisi penjelasan mengenai implementasi yang sudah dibuat, dan analisis dari aplikasi yang sudah dibangun.

**BAB V      KESIMPULAN DAN SARAN**

Bab V Kesimpulan dan Saran berisi kesimpulan dan saran yang dapat digunakan untuk mengembangkan penelitian ini di kemudian hari.