

# 1. Pendahuluan

## 1.1. Latar Belakang

Opini suatu produk memiliki peran penting sebagai informasi bagi konsumen dan produk manufaktur. Seiring dengan kemajuan teknologi, maka semakin memudahkan konsumen dalam memberikan opini terhadap produk, salah satunya melakukan *review* produk dengan memanfaatkan media elektronik dan internet. Untuk mengetahui tingkat kepuasan konsumen, produk manufaktur mengizinkan konsumen untuk *mereview* atau mengekspresikan opini mereka terhadap suatu produk. *Review* suatu produk menjadi sumber informasi yang berguna baik bagi konsumen maupun produk manufaktur [1]. Konsumen dapat menggunakan informasi tersebut sebagai pendukung pengambilan keputusan untuk membeli suatu barang. Sedangkan bagi produk manufaktur, *review* digunakan sebagai bahan evaluasi tingkat kepuasan konsumen terhadap produk, perkembangan suatu produk, maupun dalam hal pemasaran produk. Namun, dengan banyaknya jumlah *review* suatu produk, membuat sulit konsumen maupun produk manufaktur dalam mengevaluasi *review* secara cepat dan tepat. Untuk mengatasi hal tersebut dibutuhkan sebuah sistem yang dapat secara otomatis meringkas kumpulan opini dari *review* produk agar lebih efektif dan efisien.

Dalam membangun sistem peringkasan ekstraktif ini ada beberapa pendekatan *machine learning* yang dapat diterapkan. Salah satu metode yang dapat diterapkan adalah *Support Vector Regression (SVR)*. SVR merupakan model regresi dari *Support Vector Machine (SVM)* dengan mengimplementasikan *Structural Risk Minimization (SRM)* yaitu prinsip induktif untuk mendapatkan generalisasi yang baik pada dataset *training* yang terbatas [2]. Kelebihan menggunakan metode SVR ini adalah tidak bergantung pada dimensi dan vektor ruang input, sehingga pada kasus *nonlinier*, SVR dapat memetakan input ke suatu ruang dimensi yang lebih tinggi.

Pada Tugas Akhir ini, model regresi yang diimplementasikan adalah *Support Vector Regression (SVR)*. SVR digunakan untuk mengkombinasikan fitur secara otomatis dan melakukan pembobotan kalimat. Ada beberapa tahap yang dilakukan pada penelitian Tugas Akhir ini, yaitu (1) *preprocessing*, pada tahap ini dilakukan tokenisasi kalimat, *case folding*, *stopword removal*, dan *lemmatization*, (2) ekstraksi kalimat, terdiri dari tiga langkah yaitu perhitungan fitur pembobotan kalimat, skoring kalimat dan pemilihan kalimat, (3) pembentukan ringkasan dimulai dengan pengklasifikasian dan pengurutan kalimat berdasarkan kepentingan suatu kalimat.

## 1.2. Perumusan Masalah

Dengan permasalahan yang dipaparkan pada latar belakang, didapat rumusan masalah yang diangkat dalam Tugas Akhir ini adalah semakin mudahnya pengguna dalam memberikan opini terhadap suatu produk menghasilkan jumlah *review* produk

yang banyak sehingga kesulitan untuk mengambil informasi suatu produk tersebut. Salah satu dari beberapa teknik yang bisa digunakan adalah metode *Support Vector Regression* (SVR). Pada kasus ini dibutuhkan orientasi opini pada setiap kalimat, dan metode SVR dapat menangani dengan mempelajari data-data yang telah ada sebelumnya. Selain itu SVR bekerja dengan mengidentifikasi subset dari poin informatif dan menggunakannya sebagai representasi dari solusi.

### 1.3. Tujuan

Berdasarkan paparan perumusan masalah yang telah diuraikan dapat disimpulkan bahwa tujuan penelitian Tugas Akhir ini adalah:

1. Menentukan peringkasan *review* produk dengan menerapkan metode *Support Vector Regression* (SVR) dalam memprediksi kalimat.
2. Menentukan parameter optimal yang digunakan pada metode *Support Vector Regression* (SVR).
3. Menganalisis performansi sistem dengan metode *Support Vector Regression* (SVR).

### 1.4. Batasan Masalah

Adapun batasan masalah pada Tugas Akhir ini sebagai berikut:

1. *Dataset* yang digunakan merupakan dokumen sembilan *review* produk dalam bahasa Inggris yang digunakan pada paper Minqing Hu dan Bing Liu [3].
2. Metode pembobotan kalimat yang digunakan adalah perhitungan TF-IDF.
3. Fitur yang digunakan dalam proses hanya fitur eksplisit.
4. Panjang ringkasan yang dihasilkan tidak ditentukan.
5. Komposisi perbandingan data training dan testing yang digunakan yaitu 80:20.

### 1.5. Metodologi Penyelesaian

Penyelesaian Tugas Akhir ini dilakukan dengan menggunakan metodologi sebagai berikut:

1. Studi Literatur
  - a) Studi literatur berupa pencarian dasar dan referensi pada buku dan jurnal ilmiah yang berhubungan dengan *opinion summarization* dan metode *Support Vector Regression* (SVR). Tujuan dari studi literatur untuk memahami konsep dari kandidat metode yang digunakan untuk menyelesaikan masalah.
  - b) Pendalaman materi yang berhubungan dengan Tugas Akhir.
2. Pengumpulan Data

Data yang digunakan untuk penelitian merupakan *file* dengan format txt yang berisi *review* produk yang digunakan pada paper Minqing Hu dan Bing Liu [3].
3. Perancangan Sistem

Merancang sistem peringkasan dokumen, dimulai dari input, proses peringkasan hingga sistem menghasilkan output berupa ringkasan. Rancangan sistem

digambarkan dalam *flowchart* untuk memahami alur proses yang terjadi di dalamnya.

#### 4. Implementasi dan Pengujian Sistem

Implementasi dilakukan berdasarkan rancangan sistem yang sudah dibuat. Pada Tugas Akhir ini implementasi sistem dibangun dengan menggunakan bahasa pemrograman Java dengan Netbeans IDE.

#### 5. Analisis Hasil Pengujian

Melakukan analisis terhadap akurasi dari output sistem, serta pengaruh parameter dan fitur dalam menghasilkan ringkasan.

#### 6. Penyusunan Laporan Tugas Akhir

Penyusunan laporan bertujuan sebagai dokumentasi hasil penelitian dengan melampirkan dokumen pendukung pada penelitian Tugas Akhir.