# BAB 1
# PENDAHULUAN

## 1.1 Latar Belakang Masalah

*Support Vector Machine* (*SVM*) merupakan salah satu metode pembelajaran mesin bersifat *supervised* (*supervised machine learning*) yang digunakan untuk proses regresi dan klasifikasi [1]. Dari hasil beberapa penelitian yang mengkaji baik secara teori maupun implementasi [2] [3] [4], *Support Vector Machine* memiliki beberapa kelebihan, diantaranya adalah tingkat akurasi yang tinggi serta impelementasinya yang relatif mudah. *Support Vector Machine* telah banyak diimplementasikan dalam berbagai permasalahan dunia nyata (*real-world problems*), salah satunya adalah untuk peringkasan teks otomatis (*automatic text summarization*) [2].

Berdasarkan hasil penelitian yang dilakukan [1], *Support Vector Machine* memiliki beberapa kelemahan. Salah satu dari kelemahan tersebut antara lain proses prediksinya yang tidak bersifat probabilistik yang dapat mengurangi tingkat akurasi dari klasifikasi [1]. Oleh karena itu dibuatlah sebuah metode klasifikasi dan regresi yang dapat mengatasi kekurangan-kekurangan yang dimiliki oleh *Support Vector Machine*. Metode tersebut adalah *Relevance Vector Machine* (*RVM*).

*Relevance Vector Machine* adalah sebuah metode pembelajaran mesin yang diadaptasi dari *Bayesian Framework* dan memiliki bentuk model fungsi yang mirip dengan *Support Vector Machine*. Dari hasil berbagai penelitian dan perbandingan yang dilakukan menunjukkan bahwa *Relevance Vector Machine* dalam melakukan regresi dan klasifikasi memiliki tingkat akurasi yang lebih baik dengan waktu komputasi lebih cepat dari *Support Vector Machine* [1] [5], yang ditunjukkan dengan jumlah *relevant vector* (*RV*) yang dihasilkan oleh *Relevance Vector Machine* jauh lebih sedikit dari jumlah *support vector* (*SV*) yang dihasilkan oleh *Support Vector Machine.*

Penelitian kali ini mencoba untuk mengimplementasikan metode *Relevance Vector Machine* pada peringkasan teks otomatis pada dokumen tunggal (*single*

*document)*. Sama halnya dengan metode pembelajaran mesin lainnya, metode *Relevance Vector Machine* ini akan digunakan pada saat proses klasifikasi kalimat hasil *preprocessing* kalimat dalam dokumen. Kalimat dalam dokumen akan dibagi menjadi dua kelas, yaitu kelas positif dan kelas negatif. Kalimat positif berisi kalimat yang termasuk ke dalam ringkasan, sedangkan kalimat negatif tidak termasuk ke dalam ringkasan [2].

Dengan dilakukannya penelitian ini diharapkan bahwa hasil dari penelitian dapat dijadikan pendekatan baru dalam mengembangkan peringkasan teks otomastis yang sampai saat ini masih terus dikembangkan dan diteliti [6].

## 1.2 Rumusan Masalah

Berdasarkan uraian latar belakang masalah yang ada, maka perumusan masalah yang ada adalah bagaimana mengimplementasikan metode *Relevance Vector Machine* pada membuat peringkasan teks otomatis.

## 1.3 Maksud dan Tujuan

Maksud dari penelitian ini adalah untuk mengimplementasikan metode *Relevance Vector Machine* dalam pembuatan peringkasan teks otomatis. Sedangkan tujuan yang ingin dicapai dari penelitian ini adalah melihat akurasi metode *Relevance Vector Machine* dalam membuat sebuah ringkasan.

## 1.4 Batasan Masalah

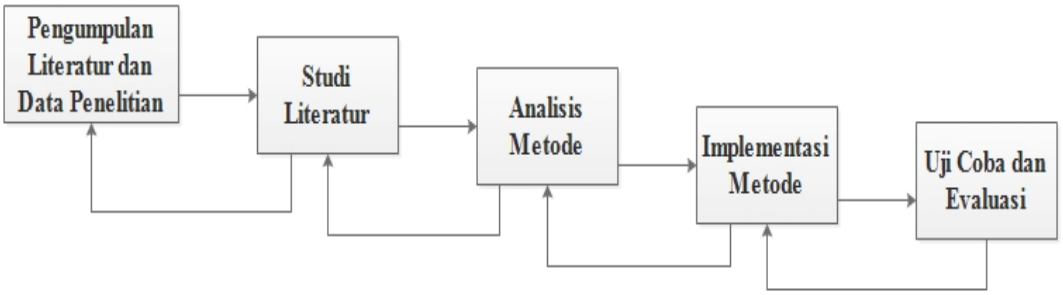
Batasan masalah merupakan pembatasan ruang lingkup pekerjaan yang akan dilakukan terhadap permasalahan yang ada. Batasan dalam penelitian ini adalah sebagai berikut.

a. Dokumen yang digunakan dalam penelitian adalah dokumen tunggal (*single document*).
b. Dokumen teks yang dijadikan bahan penelitian merupakan dokumen berbahasa Indonesia.
c. Format *file* yang digunakan adalah format *.txt*.

d. Metode peringkasan yang digunakan adalah metode ekstraksi, dimana metode ekstraksi mengambil beberapa bagian dari dokumen yang dianggap penting lalu menyusunnya kembali.
e. Hasil dari peringkas teks otomatis ini adalah berupa ringkasan.
f. Bahasa pemrograman yang akan digunakan adalah bahasa *C#*.
g. Perangkat lunak peringkasan yang akan dikembangkan berbasis *desktop*.

## 1.5 Metodologi Penelitian

Metodologi penelitian yang akan digunakan dalam penelitian ini dapat dilihat pada Gambar 1.1.




**Gambar 1.1 Metodologi Penelitian**

Adapun detail dari metodologi penelitian pada Gambar 1.1 adalah sebagai berikut.

a. Pengumpulan Literatur dan Data Penelitian

   Pengumpulan data dilakukan dengan cara mengumpulkan berbagai literatur dan referensi yang berkaitan dengan penelitian yang dilakukan, mulai dari buku, jurnal, dan artikel. Selain itu dilakukan juga pengumpulan dokumen-dokumen berbahasa Indonesia yang akan dijadikan bahan pengujian dari penelitian yang akan dilakukan.

b. Studi Literatur

   Studi literatur dilakukan dengan cara tinjauan literatur (*literature review*). Tinjauan literatur berisi uraian tentang teori, temuan dan bahan penelitian lain

yang diperoleh dari bahan acuan untuk dijadikan landasan kegiatan penelitian [7]. Tinjauan literatur dilakukan dengan cara membaca beberapa referensi-referensi pendukung penelitian yang telah dikumpulkan sebelumnya, lalu meringkas referensi-referensi tersebut sehingga membentuk pemahaman yang lebih mendalam terhadap hal-hal yang akan diteliti.

c. Analisis Metode

Dalam penelitian ini analisis dibagi menjadi beberapa bagian, yaitu analisis masalah yang diangkat, analisis data masukan, analisis tahap *preprocessing*, analisis peringkasan dengan melakukan penghitungan manual dari metode yang digunakan, serta analisis kebutuhan fungsional dan non fungsional dari perangkat lunak yang akan dibangun. Setelah analisis dilakukan, dibuat suatu perancangan model perangkat lunak yang sebelumnya telah dianalisis kebutuhan fungsionalnya, dimana perangkat lunak yang dibangun untuk mensimulasikan metode yang diterapkan.

d. Implementasi Metode

Setelah melakukan tahapan analisis metode *Relevance Vector Machine* yang akan digunakan untuk peringkasan teks otomatis, langkah selanjutnya adalah implementasi metode tersebut dengan cara membangun perangkat lunak berbasis *desktop* menggunakan bahasa pemrograman *C#*. Implementasi dilakukan berdasarkan analisis dan perancangan yang telah dilakukan sebelumnya dari mulai dari tahap *preprocessing*, tahap pelatihan, hingga ke tahap pengujian hasil setelah proses pelatihan.

e. Uji Coba dan Evaluasi

Melakukan uji coba dengan menggunakan perangkat lunak yang telah dibuat. Uji coba yang dilakukan adalah dengan melakukan proses pengujian dengan menggunakan dokumen-dokumen uji yang telah disiapkan, melihat akurasi dari setiap percobaan, serta membuat kesimpulan terhadap hasil percobaan tersebut apakah metode *Relevance Vector Machine* memiliki akurasi yang baik atau tidak.

## 1.6 Sistematika Penulisan

Sistematika penulisan laporan penelitian ini disusun untuk memberikan gambaran umum mengenai penelitian yang dilakukan. Sistematika penulisannya sebagai berikut.

BAB 1 PENDAHULUAN

Bab ini terdiri dari latar belakang masalah, perumusan masalah, maksud dan tujuan, batasan masalah, metodologi penelitian yang digunakan, serta sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini menjelaskan landasan teori yang berisi teori-teori yang berkaitan dengan penelitian, mulai dari teori peringkasan teks otomatis, *preprocessing* pada teks, pembobotan *term* menggunakan *tf-idf*, *Relevance Vector Machine*, hingga model dari analisis dan perancangan yang akan dilakukan.

BAB 3 ANALISIS KEBUTUHAN ALGORITMA

Pada bab ini dijelaskan secara rinci mulai dari analisis masalah, analisis sistem, analisis data masukan, analisis tahap *preprocessing*, serta analisis pelatihan dan klasifikasi menggunakan metode *Relevance Vector Machine*. Selain itu juga dijelaskan analisis kebutuhan fungsional dan non-fungsional serta perancangan menu, antarmuka, pesan, jaringan semantik, dan perancangan *method*.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini menjelaskan implementasi dari analisis dan perancangan sistem sebelumnya ke dalam bentuk bahasa pemrograman berbasis *desktop* yaitu *C#*. Hasil implementasi selanjutnya akan dilakukan pengujian, mengingat bahwa sistem yang dapat digunakan dengan baik harus melalui tahapan pengujian. Pengujian yang dilakukan adalah menguji akurasi hasil ringkasan teks dengan hasil ringkasan manual dan menguji peringkasan teks otomatis kepada pengguna.

# BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari keseluruhan isi laporan skripsi serta saran yang diuraikan dengan tujuan agar adanya pengembangan peringkas teks otomatis dengan mengimplementasikan metode yang lebih baik lagi untuk hasil akurasinya.