

BAB 1

PENDAHULUAN

1.1. Latar Belakang Masalah

Suatu informasi dapat diperoleh dengan menggunakan berbagai cara, salah satunya adalah bertanya pada seseorang. Namun setiap orang mempunyai keterbatasan dalam hal informasi yang dimiliki ataupun waktu yang dimiliki untuk menjawab pertanyaan. Hal ini dapat menjadi masalah bagi orang yang bertanya maupun yang ditanya, terutama bagi suatu bidang usaha yang memiliki banyak pelanggan. Masalah keterbatasan informasi dan waktu yang dihadapi seseorang dapat diatasi dengan adanya Sistem Tanya Jawab. Sistem ini dapat memberikan informasi yang relevan sesuai dengan pertanyaan yang diberikan. Misalnya seseorang ingin mengetahui cara pembayaran di suatu kampus, sistem ini akan membantu orang tersebut dengan memberikan jawaban berupa tahapan proses pembayaran.

Penelitian sebelumnya, sistem tanya jawab pada kasus front office UNIKOM telah dilakukan dengan menggunakan metode SVM [1]. Penelitian tersebut menghasilkan akurasi jawaban sebesar 39%. Salah satu faktor penyebab rendahnya akurasi jawaban terletak pada tahap klasifikasi pertanyaan. Pada tahap tersebut SVM mengklasifikasi pertanyaan dengan akurasi hanya sebesar 46%.

Oleh karena itu pada penelitian ini akan mencoba menggunakan metode klasifikasi yang berbeda. Berdasarkan penelitian yang dilakukan oleh Rafi dan Shaikh [2], metode RVM memiliki akurasi lebih baik dibandingkan SVM dengan hasil persentase f-measure untuk RVM adalah 92,33%, sedangkan untuk SVM adalah 90,29%. Selain itu, penelitian yang dilakukan oleh Bishop dan Tipping menunjukkan bahwa RVM memiliki tingkat error yang lebih rendah dibanding SVM, yaitu 9,3% untuk RVM dan 10,6% untuk SVM. Berdasarkan keunggulan

yang dimiliki RVM dibanding SVM, maka penelitian ini akan menganalisis akurasi pengklasifikasian pertanyaan dengan menggunakan RVM. Sehingga diharapkan dengan metode ini dapat menghasilkan akurasi pengklasifikasian dan jawaban yang lebih baik.

1.2. Identifikasi Masalah

Berdasarkan latar belakang yang telah disampaikan di atas, maka permasalahan yang dihadapi adalah akurasi pengklasifikasian pertanyaan dan penentuan jawaban yang masih rendah dari sistem tanya jawab.

1.3. Maksud dan Tujuan

Berikut merupakan maksud dan tujuan dari penelitian ini.

1.3.1. Maksud

Maksud penelitian ini adalah membangun sistem tanya jawab berbahasa Indonesia pada kasus *front office* UNIKOM dengan menggunakan metode RVM.

1.3.2. Tujuan

Tujuan dari penelitian ini adalah mengukur akurasi pengklasifikasian dan jawaban yang bisa didapat pada sistem tanya jawab berbahasa Indonesia dengan menggunakan metode RVM.

1.4. Batasan Masalah

Berikut ini merupakan batasan-batasan masalah dalam melakukan penelitian ini.

Data Masukan:

1. Pertanyaan yang diajukan menggunakan bahasa Indonesia.
2. Pertanyaan hanya mengenai informasi-informasi penerimaan mahasiswa baru dan autodebet.

Proses:

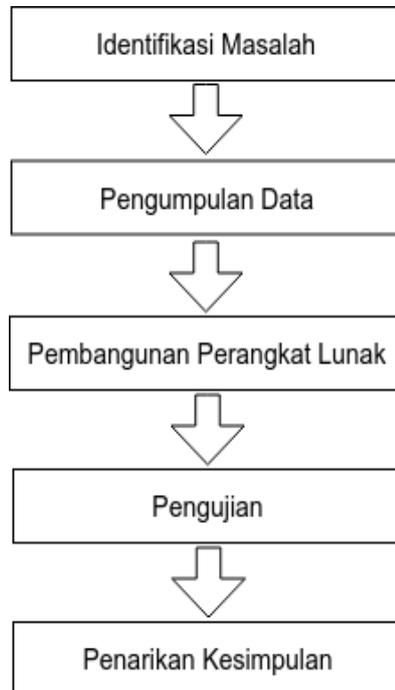
1. Tahap preprocessing yang dilakukan adalah *case folding*, *cleansing*, *tokenizing*, *stemming* dan *stopword removal*.
2. Kelas terdiri dari Waktu Autodebet, Informasi PMB, Waktu PMB, Lokasi PMB, Informasi Autodebet, Biaya PMB, dan Biaya Autodebet.

Keluaran:

1. Jawaban atas pertanyaan yang diajukan.
2. Nilai akurasi RVM
3. Nilai akurasi jawaban.

1.5. Metode Penelitian

Metode yang akan digunakan pada penelitian ini adalah metode eksperimen. Metode ini dipilih karena pengujian hipotesis akan dilakukan menggunakan serangkaian percobaan [3]. Adapun gambaran tahapan dalam penelitian ini seperti pada Gambar 1.1.



Gambar 1.1 Tahapan Metode Eksperimen

1.5.1. Identifikasi Masalah

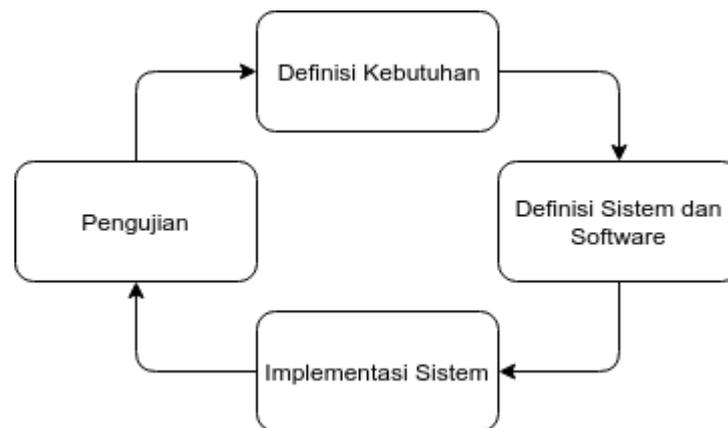
Identifikasi masalah dengan cara menganalisis penelitian sebelumnya yang sudah dilakukan dimana salah satu hasil analisis adalah rendahnya akurasi SVM pada penelitian sebelumnya.

1.5.2. Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah dengan menggunakan metode studi literatur. Studi literatur adalah cara pengumpulan data dengan menelusuri sumber-sumber tulisan yang pernah disusun sebelumnya [3]. Tulisan tersebut berupa hasil penelitian yang berhubungan dengan sistem tanya jawab dan metode RVM.

1.5.3. Pembangunan Perangkat Lunak

Metode pembangunan perangkat lunak yang digunakan adalah metode *prototyping*. Metode ini dipilih karena hasil dari pembangunan dapat dievaluasi kembali jika terjadi kesalahan [4]. Berikut adalah tahapan-tahapan pembangunan perangkat lunak menggunakan metode *prototyping*.



Gambar 1.2 Model Prototyping

1. Definisi Kebutuhan

Pada tahap ini, menetapkan seluruh kebutuhan pembangunan sistem tanya jawab [5]. Kebutuhan tersebut meliputi analisis data pertanyaan sebagai data masukan hingga klasifikasi data pertanyaan, penggunaan metode RVM sebagai klasifikasi data pertanyaan, dan penggunaan metode *cosine similarity* untuk pemberian jawaban dari sistem tanya jawab.

2. Desain Sistem dan Software

Dalam tahap ini membuat gambaran yang dikerjakan dan bagaimana tampilan dari sistem yang dibangun [5].

3. Implementasi Sistem

Tahap ini merupakan implementasi berupa perangkat lunak berdasarkan rancangan yang telah dibuat. Pembuatan perangkat lunak dipecah menjadi modul-modul kecil yang nantinya akan digabungkan dalam sistem [5].

4. Pengujian

Tahap ini dilakukan pengujian sistem terhadap modul yang dibuat, apakah sudah memenuhi fungsi yang diinginkan atau belum [5].

1.5.4. Pengujian

Pada tahap ini, sistem diuji coba diberi pertanyaan-pertanyaan. Kemudian dihitung akurasi berdasarkan jumlah jawaban benar dari seluruh pertanyaan yang diberikan.

1.5.5. Penarikan Kesimpulan

Pada tahap ini akan dilakukan analisis kesimpulan berdasarkan hasil dari pengujian sistem pada pembangunan perangkat lunak. Hasil yang diambil adalah dengan mencocokkan jawaban yang dihasilkan dari sistem tanya jawab dengan sumber data yang valid yang dijadikan sebagai tolak ukur dari akurasi sistem tanya jawab. kesimpulan dibuat menurut analisis yang sudah dilakukan.

1.6. Sistematika Penulisan

Penelitian ini disusun dengan sistematika penulisan sebagai berikut.

BAB 1 PENDAHULUAN

Bab ini menjelaskan secara garis besar mengenai latar belakang masalah, identifikasi masalah, maksud dan tujuan, batasan masalah, metode penelitian, serta sistematika penulisan dari laporan penelitian ini.

BAB 2 LANDASAN TEORI

Bab ini menjelaskan teori-teori yang berkaitan dengan sistem tanya jawab, *preprocessing (case folding, cleansing, tokenizing, stopword removal, stemming)*, Relevance Vector Machine, *cosine similarity*, akurasi, pemodelan sistem (UML), bahasa pemrograman, perangkat lunak pendukung.

BAB 3 ANALISIS DAN PERANCANGAN

Bab ini membahas tentang analisis sistem, analisis pemecahan masalah, analisis kebutuhan fungsional dan non fungsional, serta perancangan sistem yang mencakup perancangan antar muka, alur data, basis data, struktur menu, dan jaringan semantik.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi tentang hasil implementasi dari rancangan sistem yang telah dibangun sesuai dengan perancangan sistem yang telah dibuat, juga disertakan pengujian dari sistem tersebut. Pengujian akan dilakukan secara offline yakni membandingkan jawaban yang relevan terhadap pertanyaan yang diajukan oleh pengunjung atau pengguna.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan dari hasil pembangunan sistem yang telah diimplementasikan disertai juga dengan saran yang diharapkan dapat berguna untuk tahap penelitian lanjutan dari sistem ini.

