

ABSTRAK

Saran dan keluhan merupakan salah satu hal yang menjadi pertimbangan suatu organisasi demi meningkatkan pelayanan yang diberikan. Penelitian ini mengolah saran dan keluhan dari mahasiswa yang dikumpulkan melalui kuesioner pada Sistem Informasi Perguruan Tinggi XYZ selama satu tahun untuk pertimbangan dalam meningkatkan kualitas layanan Institusi. Saran dan keluhan berupa teks sehingga dibutuhkan tahapan *preprocessing* teks ulasan dari data yang didapatkan seperti proses *tokenizing*, *stemming*, dan *stopword removal*. Tujuan dari penelitian ini untuk membangun sebuah *dashboard monitoring* yang menampilkan informasi tentang saran, tahun, dan klasifikasi data teks sesuai dengan bidang perbaikan yang dibutuhkan. Metode yang digunakan yakni *sentiment analysis* dan penentuan *context*. Langkah pertama mengumpulkan data keluhan mahasiswa melalui Sistem Informasi Perguruan Tinggi XYZ yang telah dikumpulkan oleh Satuan Penjaminan Mutu (SPM) Perguruan Tinggi XYZ. Selanjutnya melakukan teknik *preprocessing data* untuk membersihkan data saran dan keluhan yang kosong. Kemudian *sentiment analysis* dilakukan untuk mengevaluasi apakah teks yang dikumpulkan memiliki *sentiment* positif, negatif atau netral. Selanjutnya, diterapkan perhitungan *Term Frequency-Inverse Document Frequency* (TF-IDF) pada masing-masing data berlabel *sentiment* positif, negatif, dan netral untuk mengekstraksi konteks dari data yang berlabel tersebut berdasarkan metode *Context-based Reasoning* (CBR). Berdasarkan CBR, konteks yang didapatkan akan dianalisis untuk dicari apa saja keluhan yang paling sering diutarakan mahasiswa, hingga objek yang paling sering disebut hingga menyebabkan opini bernilai negatif. Data hasil ini kemudian divisualisasikan untuk menjadi dasar pengambilan keputusan perbaikan fasilitas. Penggunaan *Bag of Words* pada *sentiment analysis* mendapatkan hasil yang optimal yakni dengan akurasi nilai untuk *sentiment analysis* adalah sebesar 91%. Penerapan metode SVM untuk melakukan klasifikasi pada bidang-bidang mendapatkan akurasi 89% yang mana model ini cukup optimal. Visualisasi data menggunakan Google Data Studio sebagai alat untuk membuat *dashboard* dari *dataset* tersebut yang sesuai dengan kebutuhan. Salah satu *output* pada penelitian ini berupa kumpulan kata-kata yang merepresentasikan bidang-bidang pada Institusi. Terdiri dari Bidang 1, Bidang 2, dan Bidang 3 *corpus* ini dibuat khusus untuk melakukan klasifikasi

jawaban aduan mahasiswa terhadap pelayanan yang diberikan oleh Institusi serta ditampilkan dalam *dashboard monitoring*. Hasil penelitian ini digunakan untuk rekomendasi para pemangku keputusan dalam menetapkan langkah-langkah yang akan dilaksanakan demi meningkatkan pelayanan.

Kata Kunci: *Context-based Reasoning, Bag of Words, Text Mining, Support Vector Machine, Sentiment Analysis*

ABSTRACT

Suggestions and complaints are one of the things that an organization considers in order to improve the services provided. This research processes suggestions and complaints from students collected through questionnaires on the XYZ College Information System for one year for consideration in improving the quality of Institution services. Suggestions and complaints are in the form of text so that the preprocessing stage of the review text of the data obtained is needed such as tokenising, stemming, and stopword removal. The purpose of this research is to build a monitoring dashboard that displays information about suggestions, years, and classification of text data according to the required improvement areas. The method used is sentiment analysis and context determination. The first step is to collect student complaint data through the XYZ College Information System that has been collected by the XYZ College Quality Assurance Unit (SPM). Next, perform data preprocessing techniques to clean empty suggestion and complaint data. Then sentiment analysis is performed to evaluate whether the collected text has a positive, negative or neutral sentiment. Next, Term Frequency-Inverse Document Frequency (TF-IDF) calculation is applied to each data labeled with positive, negative and neutral sentiment to extract the context of the labeled data based on Context-based Reasoning (CBR) method. Based on CBR, the context obtained will be analyzed to find out what complaints are most often expressed by students, to the objects that are most often mentioned to cause negative opinions. This result data is then visualized to become the basis for making decisions on facility improvements. The use of Bag of Words in sentiment analysis obtained optimal results, namely with an accuracy value for sentiment analysis of 91%. The application of the SVM method to classify the fields obtained an accuracy of 89%, which is quite optimal. Data visualization uses Google Data Studio as a tool to create a dashboard from the dataset as needed. One of the outputs in this research is a collection of words that represent the fields at the Institution. Consisting of Field 1, Field 2, and Field 3, this corpus is specifically made to classify the answers to student complaints about the services provided by the Institution and displayed in the monitoring

dashboard. The results of this research are used for recommendations for decision makers in determining steps to be taken to improve services.

Keywords: Context-based Reasoning, Bag of Words, Text Mining, Support Vector Machine, Sentiment Analysis