

Pemanfaatan Education Data Mining (EDM) dalam pengelolaan data survey umpan balik mahasiswa terkait peningkatan mutu perguruan tinggi: studi kasus Universitas XYZ = Utilizing Education Data Mining (EDM) to improve the quality of higher education: case study of XYZ University

Calandra Alencia Haryani, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20497777&lokasi=lokal>

Abstrak

Universitas XYZ merupakan salah satu PTS dalam bentuk universitas yang berlokasi di Tangerang, yang memiliki kewajiban untuk menjamin dan memberikan pendidikan bermutu kepada mahasiswa selaku salah satu pemangku kepentingan yang memiliki dampak secara langsung pada mutu sebuah universitas. LP2MP bertugas untuk menyelenggarakan pengukuran dan pelaksanaan survei setiap semester dalam bentuk survei umpan balik. Hasil survei tersebut dapat dijadikan sebagai pedoman untuk perbaikan yang berkesinambungan untuk penyelenggaraan penjaminan mutu Dikti dan pengelolaan Universitas XYZ. Namun, pengolahan dan pengukuran data survei secara konvensional tidak cukup untuk mengeksplorasi informasi tersembunyi dari survei. EDM digunakan pada penelitian ini untuk mengolah dan menganalisa data dari Universitas XYZ berupa survei bagian Open Ended Question (OEQ) yang terdiri dari Student Feedback Questionnaire (SFQ), Facility Satisfaction Questionnaire (FSQ), dan Graduate Feedback Questionnaire (GFQ).

Tujuan dari penelitian ini adalah untuk mendukung pengambilan keputusan dalam mengambil tindakan proaktif terhadap perbaikan mutu Universitas XYZ. Penelitian ini melakukan klasifikasi label aspek, sentimen analisis, dan tren topik survei SFQ, FSQ, dan GFQ pada bagian OEQ. Klasifikasi label aspek Multi Class survei SFQ memilih model klasifikasi terbaik dengan membandingkan hasil evaluasi accuracy, precision, recall, dan F1-Score terhadap setiap kombinasi fitur dan perbandingan empat algoritma klasifikasi yaitu Decision Tree (DT), Naive Bayes (NB), K-Nearest Neighbor (KNN), dan Support Vector Machine (SVM). Klasifikasi label aspek multi label survei FSQ dan GFQ memilih model klasifikasi terbaik dengan membandingkan hasil evaluasi tiga jenis library multilabel dari SciKit-Learn, yaitu Binary Relevance (BL), Label Power Set (LPS), dan Classifier Chain (CC) terhadap setiap kombinasi fitur dan empat algoritma klasifikasi tersebut.

Hasil dari penelitian ini adalah teknik klasifikasi menggunakan kombinasi fitur TFIDF, Unigram, dan Bigram dengan algoritma SVM merupakan model klasifikasi terbaik untuk pelabelan aspek survei SFQ. Teknik klasifikasi menggunakan kombinasi fitur TFIDF, Unigram, dan Bigram dengan algoritma SVM dan library Multi Label CC merupakan model klasifikasi terbaik untuk pelabelan aspek survei FSQ. Teknik klasifikasi menggunakan kombinasi fitur Count Vectorizer, Unigram, dan Bigram dengan algoritma NB dan library Multi Label BR merupakan model klasifikasi terbaik untuk pelabelan aspek survei GFQ. Selain itu, algoritma SentiStrenghtID digunakan untuk mendapatkan sentimen dan algoritma LDA digunakan untuk mendapatkan tren topik tahunan pada setiap label aspek survei SFQ, FSQ, dan GFQ.

.....XYZ University is one of the universities in the form of universities located in Tangerang, which has an obligation to guarantee and provide quality education to students as one of the stakeholders that has a direct impact on the quality of a university. LP2MP is tasked to carrying out measurements and implementation of

feedback every semester in the form of surveys as one part of quality control directly to stakeholders. The results of the surveys can be used as a guideline for continuous improvement in the implementation of Dikti quality assurance and management of XYZ University. However, conventional processing and measurement of feedback data are not enough to explore hidden information from surveys data. EDM was used in this research to process and analyze data from XYZ University in the form of Student Feedback Questionnaire (SFQ), Facility Satisfaction Questionnaire (FSQ), and the Graduate Feedback Questionnaire (GFQ) in the Open-Ended Question (OEQ) section.

The purpose of the research is to support decision making in taking proactive actions towards improvement for self-evaluation and quality of XYZ University. This research carried out label aspect classification, analytical sentiment, and trends in the survey topics SFQ, FSQ, and GFQ in the OEQ section. Multi-class aspect label classification SFQ will choose the best classification model by comparing the results of the evaluation of accuracy, precision, recall, and F1-score for each feature combination and comparison of four classification algorithms namely Decision Tree (DT), Naive Bayes (NB), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM). The classification of the multi-label aspects of the FSQ and GFQ survey labels will have the best classification model by comparing the evaluation results of three multilabel library types from SciKit-Learn, namely Binary Relevance (BR), Label Power Set (LPS), and Classifier Chain (CC) to each combination of features and four classification algorithms.

The results of this research are Classification Techniques using a combination of features of TFIDF, Unigram, and Bigram with the SVM algorithm which is the best Multi Class classification model for labeling aspects of the SFQ survey. Classification techniques use a combination of TFIDF, Unigram, and Bigram features with the SVM algorithm and the Multi-Label library CC is the best Multi-Label classification model for labeling aspects of the FSQ survey. Classification techniques using a combination of Count Vectorizer, Unigram, and Bigram features with the NB algorithm and the Multi-Label library BR are the best Multi-Label classification models for labeling aspects of the GFQ survey. In addition, the SentiStrengthID algorithm can be used to get sentiments and the LDA algorithm can be used to get annual topic trends on each survey aspect label SFQ, FSQ, and GFQ.