

ABSTRAK

Pendidikan jarak jauh merupakan salah satu usaha yang dalam upaya pemerataan pendidikan di Indonesia. Salah satu bentuk pendidikan jarak jauh yang dinilai dapat mendukung program tersebut adalah *Massive Open Online Learning* (MOOC). Namun, hingga saat ini MOOC masih mengalami masalah pada rendahnya tingkat retensi, yaitu berkisar 5-10%. Hal ini terjadi karena pembelajaran online sangat bergantung pada faktor afektif yang merupakan sisi emosi penggunanya. Dalam penelitian ini, dilakukan penelitian terkait model klasifikasi tingkat retensi terkait faktor afektif untuk pengguna MOOC di Indonesia. Penelitian ini terdiri atas dua tahap. Tahap pertama adalah tahap untuk mengidentifikasi faktor afektif yang secara signifikan mempengaruhi tingkat retensi. Tahap kedua membangun model retensi MOOC berdasar faktor yang teridentifikasi pada tahap pertama.

Berdasarkan studi literatur terhadap 89 jurnal bereputasi, ditemukan beberapa faktor afektif yang mempengaruhi retensi seseorang selama menggunakan MOOC, khususnya di negara berkembang. Penentuan faktor yang signifikan mempengaruhi retensi MOOC di Indonesia dianalisis menggunakan SmartPLS berdasarkan data yang dikumpulkan dari 141 responden. Pada hasil uji statistik ditemukan bahwa *perceived ease of use*, *social influence*, *self-efficacy*, *power distance* dan sumber daya memiliki pengaruh yang signifikan terhadap retensi seseorang dalam penggunaan MOOC. Uji statistik menunjukkan bahwa data yang diukur memiliki kemampuan klasifikasi. Tahap kedua dilakukan dengan membangun model klasifikasi dibangun menggunakan metode *machine learning Support Vector Machine* (SVM), yang memiliki kemampuan baik dengan jumlah data yang kecil, dan telah banyak digunakan pada pelbagai penelitian terkait model klasifikasi pada MOOC. Data yang digunakan dalam pemodelan memiliki ketimpangan antara tingkat lulus dan *dropout*. Ketidakseimbangan ini dapat menyebabkan model klasifikasi yang bias dan lebih memihak pada kelas mayoritas (dalam hal ini, tingkat kelulusan). Untuk menyeimbangkan data, maka dilakukan *oversampling* sebelum membangun model klasifikasi. Dari hasil pengujian, *oversampling* menggunakan metode *Synthetic Minority Over-sampling Technique* (SMOTE)-*weighted N* dengan *chi-square* menunjukkan tingkat akurasi yang paling baik, yaitu 94,77%, *recall* 95,72%, *precision* 95,3% dan *F1-measure* 95,36%. Hasil ini menunjukkan bahwa model klasifikasi yang dibangun akurat, meskipun terjadi ketidakseimbangan data pada awal pemrosesan. Uji coba selanjutnya dilakukan dengan menerapkan faktor yang ditemukan menjadi kursus dengan fitur yang sesuai, dan terdapat peningkatan retensi sebesar 7% berdasarkan dua kelompok uji coba tersebut. Temuan ini menggarisbawahi bahwa model klasifikasi menggunakan faktor afektif ini memiliki pengaruh signifikan pada proses pembelajaran MOOC. Dengan model MOOC yang tepat, sisi afektif pengguna selama pembelajaran daring dapat dikelola secara efektif, yang berpotensi menghasilkan tingkat retensi yang lebih tinggi.

Kata kunci—MOOC; retensi; Indonesia; Hofstede; afektif; SMOTE; SVM;

ABSTRACT

Distance education is one of the efforts to equalize education in Indonesia. One form of distance education that is considered to be able to support the program is Massive Open Online Learning (MOOC). However, until now, MOOC still experiences problems with low retention rates ranging from 5-10%. This happens because online learning is highly dependent on affective factors, which are the emotional side of its users. This study conducted research on the retention rate prediction model related to affective factors for MOOC users in Indonesia. This study consists of two stages. The first stage is the stage to identify affective factors that significantly affect the retention rate. The second stage builds a MOOC retention model based on the factors identified in the first stage.

Based on a literature study of 89 reputable journals, several factors affect a person's retention while using MOOCs, especially in developing countries. Determining factors significantly affecting MOOC retention in Indonesia was analyzed using SmartPLS based on data collected from 141 respondents. The statistical test results found that perceived ease of use, social influence, self-efficacy, power distance, and resources significantly influence a person's retention in using MOOCs. Statistical tests show that the measured data has predictive ability. The second stage is carried out by building a prediction model using the Support Vector Machine (SVM) machine learning method, which has good capabilities with a small amount of data and has been widely used in various studies related to prediction models in MOOC. The data used in the modeling has an imbalance between the pass and dropout rates. This imbalance can cause a biased prediction model and favor the majority class (in this case, the graduation rate). Oversampling is carried out before building a prediction model to balance the data. From the test results, oversampling using the Synthetic Minority Over-sampling Technique (SMOTE)-weighted N method with chi-square shows the best level of accuracy, which is 94.77%, recall 95.72%, precision 95.3%, and F1-measure 95.36%. These results indicate that the prediction model built is accurate, even though there is an imbalance in the data at the beginning of processing. Further testing was conducted by implementing the discovered factors into courses with appropriate features, and there was a 7% increase in retention based on the two trial groups. This finding underlines that the prediction model using these affective factors significantly influences the MOOC learning process. With the right MOOC model, the affective side of users during online learning can be managed effectively, potentially resulting in higher retention rates.

Keywords— MOOC; retention; Indonesia; Hofstede; affective; SMOTE; SVM