

ABSTRACT

Students with learning disabilities need to be identified as early as possible to obtain appropriate learning methods and tools. One way to identify learning disabilities is to make predictions based on information that is judged on the basis of the assessment of physical and mental characteristics. But the fact that students who have learning difficulties have a small ratio compared to normal students so that obtained comparison of unbalanced data between students who have learning disabilities with normal students. This problem of imbalances can lead to low predictive performance of minority class data because the classification algorithm is more focused on the majority class, so in the research applied a class balancing technique that aims to overcome the problems in the classification data of students with learning disabilities.

In this study the data used are dataset of students with learning disabilities consisting of 208 instances with 12 attributes that have differences in minority class with a significant majority. Classification is done using integrated algorithm Correlated Naïve Bayes Classifier (C-NBC) with the Synthetic Minority Oversampling Techique (SMOTE) technique to balance the number of minority classes against the majority class. Analysis of classification performance results is done using accuracy, sensitivity and specificity.

The results of the research resulted the accuracy, sensitivity, and specificity obtained from the C-NBC algorithm classification of 94.15%, 100% and 95.61% respectively. Meanwhile, for the results obtained from the combination of C-NBC algorithm with SMOTE were respectively 94.63%, 86.51% and 97.12%. So from the results that have been obtained shows if the performance based on the accuracy and specificity of the classification of C-NBC combination with SMOTE better than the classification of C-NBC

Keywords: class imbalance, C-NBC, SMOTE, students with learning disabilities

INTISARI

Siswa berkesulitan belajar perlu diidentifikasi sedini mungkin untuk mendapatkan metode dan sarana pembelajaran yang sesuai. Salah satu cara untuk mengidentifikasi kesulitan belajar adalah dengan melakukan prediksi berdasarkan informasi yang diperoleh berdasar penilaian ciri-ciri fisik dan mental. Namun faktanya siswa yang mengalami kesulitan belajar memiliki rasio yang sedikit dibandingkan siswa normal sehingga diperoleh perbandingan data yang tidak seimbang antara data siswa yang mengalami kesulitan belajar dengan siswa normal. Masalah ketidakseimbangan ini dapat menyebabkan rendahnya performa prediksi terhadap data kelas minoritas karena algoritme klasifikasi lebih fokus terhadap kelas mayoritas, sehingga dalam penelitian diterapkan teknik penyeimbangan kelas yang bertujuan untuk mengatasi permasalahan pada klasifikasi data siswa berkesulitan belajar.

Dalam penelitian ini data yang digunakan adalah *dataset* siswa berkesulitan belajar yang terdiri dari 208 *instance* dengan 12 atribut yang memiliki perbedaan perbandingan kelas minoritas dengan kelas mayoritas yang signifikan yaitu 1 berbanding 5. Klasifikasi dilakukan menggunakan algoritme *Correlated Naïve Bayes Classifier* (C-NBC) yang diintegrasikan dengan teknik *Synthetic Minority Oversampling Technique* (SMOTE) untuk menyeimbangkan jumlah kelas minoritas terhadap kelas mayoritas. Analisis hasil kinerja klasifikasi dilakukan menggunakan nilai akurasi, sensitivitas dan *specificity*.

Hasil penelitian menghasilkan kinerja akurasi, sensitivitas, dan *specificity* yang diperoleh dari klasifikasi algoritme C-NBC masing-masing sebesar 94.15%, 100% dan 95,61%. Sementara itu, untuk hasil yang diperoleh dari kombinasi algoritma C-NBC dengan SMOTE masing-masing sebesar 94,63%, 86,51% dan 97,12%. Sehingga dari hasil yang telah diperoleh tersebut menunjukkan jika kinerja berdasarkan akurasi dan *specificity* dari klasifikasi kombinasi C-NBC dengan SMOTE lebih baik dibandingkan klasifikasi C-NBC.

Kata kunci – ketidakseimbangan kelas, C-NBC, SMOTE, siswa berkesulitan belajar