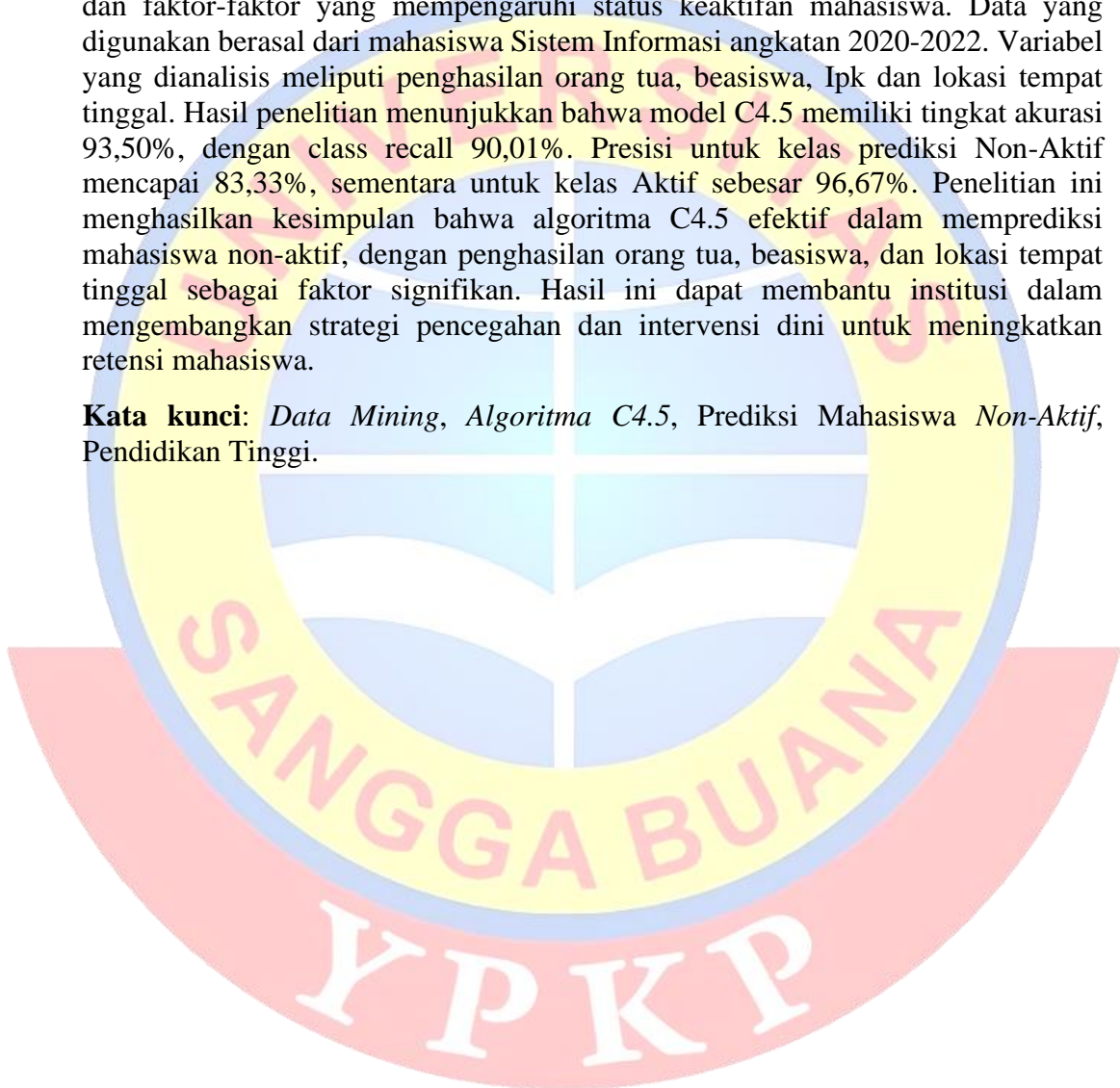


## ABSTRAK

Pendidikan tinggi memiliki peran penting dalam pembangunan sumber daya manusia dan kemajuan bangsa. Namun, masalah putus kuliah atau mahasiswa non-aktif masih menjadi tantangan serius bagi perguruan tinggi di Indonesia. Penelitian ini bertujuan untuk menganalisis dan memprediksi mahasiswa non-aktif di Universitas Sangga Buana menggunakan algoritma C4.5. Metode data mining dengan algoritma C4.5 dipilih karena kemampuannya dalam mengidentifikasi pola dan faktor-faktor yang mempengaruhi status keaktifan mahasiswa. Data yang digunakan berasal dari mahasiswa Sistem Informasi angkatan 2020-2022. Variabel yang dianalisis meliputi penghasilan orang tua, beasiswa, Ipk dan lokasi tempat tinggal. Hasil penelitian menunjukkan bahwa model C4.5 memiliki tingkat akurasi 93,50%, dengan class recall 90,01%. Presisi untuk kelas prediksi Non-Aktif mencapai 83,33%, sementara untuk kelas Aktif sebesar 96,67%. Penelitian ini menghasilkan kesimpulan bahwa algoritma C4.5 efektif dalam memprediksi mahasiswa non-aktif, dengan penghasilan orang tua, beasiswa, dan lokasi tempat tinggal sebagai faktor signifikan. Hasil ini dapat membantu institusi dalam mengembangkan strategi pencegahan dan intervensi dini untuk meningkatkan retensi mahasiswa.

**Kata kunci:** *Data Mining, Algoritma C4.5, Prediksi Mahasiswa Non-Aktif, Pendidikan Tinggi.*



## **ABTRACT**

*Higher education plays an important role in human resource development and national progress. However, the problem of dropping out of college or inactive students is still a serious challenge for universities in Indonesia. This study aims to analyze and predict inactive students at Sangga Buana University using the C4.5 algorithm. The data mining method with the C4.5 algorithm was chosen because of its ability to identify patterns and factors that affect student activeness status. The data used comes from the 2020-2022 class of Information Systems students. The variables analyzed include parental income, scholarships, ipk and location of residence. The results showed that the C4.5 model has an accuracy rate of 93.50%, with a class recall of 90.01%. Precision for the Non-Active prediction class reached 83.33%, while for the Active class it was 96.67% This study concludes that the C4.5 algorithm is effective for predicting inactive students, with parental income, scholarships, and location of residence as significant factors. These results can assist institutions in developing prevention and early intervention strategies to improve student retention.*

**Keywords:** *Data Mining, C4.5 Algorithm, Inactive Student Prediction, Higher Education.*

