

## **ABSTRACT**

Jerry Aivanca Pattikawa (00000022546)

### **CUSTOMER CHURN ANALYTICS: CASE ON UNIVERSITAS PELITA HARAPAN USING PDDIKTI DATASET**

(xiv + 54 pages: 26 figures; 8 tables, 1 appendix)

Customer is the main objective for a business' operational continuity. For a university, the students are the customers. The behaviour of a student leaving the university is called churn. Churn analysis is done to understand the student churn rate at Universitas Pelita Harapan (UPH) and the factors that might contribute to student churn.

Two models were used in this research which are Random Forest and Extreme Gradient Boosting (XGBoost). From analysis results for Random Forest were Mean Squared Error (MSE) of 21.868, Root Mean Squared Error (RMSE) of 4.676, Mean Absolute Error (MAE) of 1.872 and R-squared ( $R^2$ ) of 0.934 and the results for Extreme Gradient Boosting (XGBoost) were Mean Squared Error (MSE) of 234.352, Root Mean Squared Error (RMSE) of 15.309, Mean Absolute Error (MAE) of 7.132 and R-squared ( $R^2$ ) of 0.291. From these results, the Random Forest model is the best out of the two models. Being the best model out of the two, a confusion matrix and ROC analysis was done on the Random Forest Model. The results were the Random Forest Model had the accuracy level of 98.95%, precision level of 99.24%, recall level of 98.79% and AUC score of 0.98.

Further Analysis was done and it was found that the majors with the highest student churn rate was the majors with the worst accreditation which means accreditation had an effect on student churning.

Reference 14 (2010 – 2021)

## ABSTRAK

Jerry Aivanca Pattikawa (00000022546)

### CUSTOMER CHURN ANALYTICS: CASE ON UNIVERSITAS PELITA HARAPAN USING PDDIKTI DATASET

(xiv + 54 halaman: 26 gambar; 8 tabel, 1 lampiran)

Pelanggan adalah tujuan utama kelangsungan operasional suatu bisnis. Bagi sebuah universitas, mahasiswa adalah pelanggannya. Perilaku seorang mahasiswa yang keluar dari universitas disebut churn. Analisis churn dilakukan untuk mengetahui tingkat churn mahasiswa di Universitas Pelita Harapan (UPH) dan faktor-faktor yang mungkin berkontribusi terhadap churn mahasiswa.

Dua model yang digunakan dalam penelitian ini adalah *Random Forest* dan *Extreme Gradient Boosting* (XGBoost). Dari hasil analisa untuk *Random Forest* didapatkan *Mean Squared Error* (MSE) sebesar 21.868, *Root Mean Squared Error* (RMSE) sebesar 4.676, *Mean Absolute Error* (MAE) sebesar 1.872 dan *R-squared* ( $R^2$ ) sebesar 0.934 dan hasil untuk *Extreme Gradient Boosting* (XGBoost) adalah *Mean Squared Error* (MSE) sebesar 234.352, *Root Mean Squared Error* (RMSE) sebesar 15.309, *Mean Absolute Error* (MAE) sebesar 7.132 dan *R-squared* ( $R^2$ ) sebesar 0.291. Dari hasil tersebut, model *Random Forest* merupakan yang terbaik dari kedua model tersebut. Menjadi model terbaik dari keduanya, matriks kebingungan dan analisis ROC dilakukan pada Model *Random Forest*. Hasilnya adalah Model *Random Forest* memiliki tingkat akurasi 98,95%, tingkat *precision* 99,24%, tingkat *recall* 98,79% dan skor AUC 0,98.

Analisis lebih lanjut dilakukan dan ditemukan bahwa jurusan dengan tingkat churn siswa tertinggi adalah jurusan dengan akreditasi terburuk yang berarti akreditasi berpengaruh terhadap churn mahasiswa.

Referensi 14 (2010 – 2021)