

ABSTRAKSI

ANALISIS PERBANDINGAN KINERJA ALGORITMA C4.5, NAÏVE BAYES CLASSIFIER, DAN SUPPORT VECTOR MACHINE UNTUK SISTEM DETEKSI KEJAHATAN KARTU KREDIT

Oleh

Klaudius Liasta

15/378064/PA/16539

Kartu kredit merupakan salah satu pilihan pembayaran tanpa uang tunai yang paling banyak digunakan di seluruh dunia. Peningkatan pengguna kartu kredit sebagai alat pembayaran harus diimbangi dengan peningkatan dari sistem keamanannya. Permasalahan ini menjadi fokus utama bagi penyedia layanan kartu kredit agar pengguna tetap merasa aman dalam bertransaksi. Terdapat dua mekanisme dalam menghindari kejahatan pada kartu kredit, sistem pencegahan kejahatan dan sistem deteksi kejahatan.

Pada penelitian dilakukan analisa perbandingan menggunakan 3 algoritma klasifikasi, yaitu *Decision Tree C4.5*, *Naïve Bayes Classifier*, dan *Support Vector Machine*. Kemudian dilakukan juga ujicoba *undersampling* untuk mengetahui dampak *undersampling* pada dataset dengan ketidakseimbangan data terhadap sistem deteksi kejahatan pada kartu kredit. Hasil klasifikasi yang akan digunakan sebagai pembanding berupa nilai *accuracy*, *precision*, *recall* dan *f-1 score*.

Hasil dari penelitian ini adalah pada data dengan kondisi tidak seimbang, SVM menjadi klasifikasi terbaik dengan nilai rata-rata *recall* sebesar 75,23%, *f1-score* sebesar 82,14%, *precision* sebesar 90,45%, *accuracy* sebesar 99,94%, dan dengan waktu yang dibutuhkan 54,86 detik. Sedangkan pada data dengan kondisi seimbang karena *undersampling*, GNB menjadi klasifikasi terbaik dengan nilai rata-rata *recall* sebesar 52,17%, *f1-score* sebesar 55,11%, *precision* sebesar 58,40%, *accuracy* sebesar 91,56%, dan dengan waktu yang dibutuhkan 0,01 detik.

Kata Kunci: Sistem Deteksi Kejahatan, Kartu Kredit, Support Vector Machine, Gaussian Naïve Bayes, Decision Tree C4.5, Undersampling

ABSTRACT

COMPARATIVE ANALYSIS OF C4.5 ALGORITHM, NAÏVE BAYES CLASSIFIER, AND SUPPORT VECTOR MACHINE FOR CREDIT CARD FRAUD DETECTION SYSTEM

By

Klaudius Liasta

15/378064/PA/16539

Credit cards are one of the most widely used cashless payment options worldwide. The increase in credit card users as a means of payment must be balanced with improvements from the security system. This problem is the main focus of credit card service providers so that users continue to feel safe in the transaction. There are two mechanisms in avoiding crime on credit cards, crime prevention systems, and crime detection systems.

The purpose of the research is analyzing the best classification for the fraud detection systems using 3 classification algorithms, such as Decision Tree C4.5, Naïve Bayes Classifier, and Support Vector Machine. Then do comparison of the result on a dataset after undersampling. The classification results that will be used as a comparison consist of accuracy, precision, memory and f-1 scores.

The result of this study are SVM is a good choice to handle imbalance with average value of *recall* 75,23%, *f1-score* 82,14%, *precision* 90,45%, *accuracy* 99,94%, and consuming 54 seconds,. Otherwise, on *dataset* after *undersampling*, *GNB* be the best methode with average value of *recall* 52,17%, *f1-score* 55,11%, *precision* 58,40%, *accuracy* 91,56%, and consuming 0,01 seconds.

Keywords: Fraud Detection System, Credit Card, Support Vector Machine, Gaussian Naïve Bayes, Decision Tree C4.5, Undersampling