

INTISARI

METODE HYBRID KOMBINASI DARI MODIFIED K-PROTOTYPES DAN C5.0 UNTUK CREDIT SCORING

Oleh

Gilang Ramadhan
12/330922/PA/14405

Credit scoring merupakan salah satu teknik manajemen risiko untuk meminimalisasi terjadinya gagal bayar. Dengan menggunakan pemodelan statistika, *credit scoring* dapat dikerjakan menggunakan dua metode, yaitu *unsupervised learning* dan *supervised learning*. Perbedaan antara kedua metode tersebut terletak pada pembentukan model dan pengelompokan objek. Pada *supervised learning* terdapat proses pelatihan data dalam membentuk modelnya dan data akan diklasifikasi terhadap kelas yang telah ditentukan sebelumnya, sedangkan pada *unsupervised learning* tidak. Pada tugas akhir ini, akan digunakan kombinasi antara kedua metode tersebut yang selanjutnya dikenal dengan metode *hybrid*. Pada tahap pertama, debitur akan dikelompokkan menurut kesamaan karakteristik kedalam beberapa klaster dengan menggunakan analisis klaster *modified k-prototypes*. Metode tersebut merupakan salah satu metode *unsupervised learning*. Jumlah klaster optimal akan ditentukan menggunakan *silhouette width* dengan membandingkan beberapa jumlah klaster. Pada tahap kedua, pohon keputusan akan dibangun untuk masing-masing klaster dengan menggunakan algoritma C5.0. Algoritma tersebut merupakan salah satu metode *supervised learning*. Pada akhirnya, akan terbentuk aturan untuk masing-masing klaster dalam mengklasifikasikan debitur baik dan debitur buruk.

Kata Kunci : *Credit scoring*, analisis klaster, *modified k-prototypes*, pohon keputusan, C5.0, *silhouette width*.

ABSTRACT

HYBRID METHOD COMBINATION OF MODIFIED K-PROTOTYPES AND C5.0 FOR CREDIT SCORING

by

Gilang Ramadhan
12/330922/PA/14405

Credit scoring is one of risk management analysis method for minimizing default. Using statistical modeling, credit scoring can be used in two ways: unsupervised learning and supervised learning. The difference between these two methods are in forming the model and grouping the object. In supervised learning there is learning process in forming models and objects will be classified into pre-classified class, while in unsupervised learning doesn't need that all. In this research, will use a combination of both method, called hybrid method. In the first step, by using modified k-prototypes which is one of unsupervised learning method, debtors are segmented into clusters with similar features. The optimal number cluster is defined using silhouette width, by comparing some cluster number. In the second step, by using C5.0 algorithm which is one of supervised learning method, decision trees are built for each cluster to define classification rules for good debtor and bad debtor.

Keywords : Credit scoring, cluster analysis, modified k-prototypes, decision tree, C5.0, silhouette width.