Pemodelan credit scoring untuk kredit UMKM dengan menggunakan metode klasifikasi data mining: Studi kasus PT Pegadaian (Persero) = Credit scoring Modeling for MSME Credit using data mining classification method: case study of PT Pegadaian (Persero)

Wike Ulfiani Aresa, author

Abstrak

Saat ini credit scoring calon nasabah produk Kreasi Pegadaian masih menggunakan scorecard konvensional berupa pembobotan pertanyaan. Model credit scoring tersebut dibangun berdasarkan pengalaman pakar (expert scorecard) dan kemungkinan ada unsur subjektivitas dalam penilaian kelayakan kredit. Untuk mengatasi masalah tersebut, penelitian ini membangun model credit scoring dengan pendekatan data mining menggunakan data riwayat kredit nasabah produk Kreasi (data driven scoring) menggunakan algoritma klasifikasi, diantaranya: support vector machine (SVM), naïve bayes, decision tree dan neural network. Pengembangan model dilakukan dengan menggunakan metodologi CRISP-DM (The Cross Industry Standard Process for Data Mining). Model dibangun dengan kriteria tanpa penggunaan feature selection dan dengan feature selection. Teknik SMOTE (Synthetic Minority Over Sampling Technique) dan Oversampling dipilih untuk menyeimbangkan class data. Dari hasil evaluasi kinerja model menunjukan model SVM dengan feature selection dan penyeimbangan class menggunakan teknik Oversampling dipilih sebagai model dengan kinerja terbaik.

.....Currently, the credit worthiness of Pegadaian prospective customers still uses a conventional scorecard in the form of weighting questions. The model is built based on expert experience which is called expert scorecard. There might be an element of subjectivity in credit assessment. To resolve that problem, in this research data mining classification techniques are used to build credit scoring models. There are four classification algorithms, namely SVM (Support Vector Machine), Naïve Bayes, Decision Tree and Neural Network as a classification algorithm. Modelling uses the historical customer credit data of Pegadaian Kreasi product. CRISP-DM (The Cross Industry Standard Process for Data Mining) is used as a development methodology. Modeling is done with two criteria, by considering the use of feature selection and without feature selection. The SMOTE (Synthetic Minority Over Sampling Technique) and Oversampling techniques are chosen to balance the class data. The result of this research shows the SVM model with feature selection and data balancing using the Oversampling technique was chosen as the model with the best performance.