

INTISARI

EVALUASI PERFORMA MODEL GABUNGAN *DEEP NEURAL NETWORK* UNTUK FEATURE ENGINEERING DAN EXTREME GRADIENT BOOSTING UNTUK KLASIFIKASI BINER CREDIT SCORING

Oleh

Salsadila Aulia Fauzi

20/462330/PA/20302

Dengan semakin berkembangnya industri keuangan dan konsep konsumsi masal, bisnis kredit telah tumbuh pesat dan lembaga keuangan menghadapi tantangan yang semakin berat, salah satunya kredit macet. Adanya kredit bermasalah tentunya akan berpengaruh terhadap kesehatan dan keuntungan bank. Dalam proses ini, adanya penilaian kredit (*credit scoring*) yang merupakan teknologi pemodelan klasifikasi biner, cukup penting untuk memodelkan potensi risiko pemohon kredit. Dari banyaknya penelitian, beberapa menyebutkan bahwa model *Extreme Gradient Boosting* (XGBoost) dan *Deep Neural Network* (DNN) merupakan metode yang efektif dalam memodelkan skor kredit. DNN memiliki kemampuan untuk menangani *input* data dalam volume besar dan mengatasi nonlinearitas yang sering ditemukan dalam data kredit, serta dapat memproses informasi tersebut dengan baik. Di sisi lain, XGBoost dikenal sebagai metode klasifikasi yang unggul dalam memodelkan klasifikasi kredit, karena kinerjanya yang tinggi. Berdasarkan hal ini, diusulkan sebuah model gabungan DNN-XGBoost yang menggabungkan kekuatan kedua pendekatan tersebut. DNN akan digunakan untuk menyerap dan mengekstraksi fitur penting dari data yang besar dan nonlinear, sementara XGBoost akan memanfaatkan fitur-fitur ini untuk melakukan klasifikasi kredit. Proses penggabungan ini dimulai dari prapemrosesan data, dilanjutkan dengan ekstraksi fitur oleh DNN, dan diakhiri dengan klasifikasi oleh XGBoost. Hasil penelitian ini menunjukkan bahwa pada ketiga dataset yang digunakan, yaitu HMEQ, GERMAN, JAPANESE menunjukkan performa akurasi yang lebih baik dibandingkan dengan single model seperti *Decision Tree*, *Random Forest*, *Logistic Regression*, *Support Vector Machine*, *AdaBoost*, *Gradient Bosting Machine*, dan XGBoost itu sendiri. Dataset HMEQ dengan akurasi tertingginya yaitu 98%, German 78%, dan Japanese 90%. Model klasifikasi *credit scoring* dengan algoritma DNN-XGBoost menggunakan dataset *credit scoring* pada skripsi ini menunjukkan kinerja yang lebih baik dibandingkan beberapa single model.

Kata kunci: *Credit Scoring*, Klasifikasi Biner, Model Gabungan, Ekstraksi Fitur, *Deep Neural Network*, DNN, *Extreme Gradient Boosting*, XGBoost.

ABSTRAK

PERFORMANCE EVALUATION OF COMBINED MODEL DEEP NEURAL NETWORK FOR FEATURE ENGINEERING AND EXTREME GRADIENT BOOSTING FOR BINARY CLASSIFICATION IN CREDIT SCORING

By

Salsadila Aulia Fauzi

20/462330/PA/20302

With the rapid development of the financial industry and the concept of mass consumption, the *credit* business has grown significantly, and financial institutions face increasing challenges, including non-performing loans. Problematic loans can impact the health and profitability of banks. In this case, *credit scoring*, which is a binary classification modeling technology, is quite important to model the potential risk of *credit* applicants. From many studies, some mentioned that Extreme Gradient Boosting (XGBoost) and *Deep Neural Network* (DNN) models are effective *methods* in modeling *credit* scores. DNN excels in handling large volumes of *input* data and addressing the nonlinearity often found in *credit* data, processing this information effectively. Meanwhile, XGBoost is recognized for its superior performance in *credit* classification modeling. Based on this, a combined DNN-XGBoost model is proposed that combines the strengths of both approaches. DNN will be used to absorb and extract essential features from large and nonlinear data, while XGBoost will utilize these features for *credit* classification. The integration process begins with data *preprocessing*, followed by feature extraction by DNN, and concludes with classification by XGBoost. The results of this study show that the three datasets used, HMEQ, GERMAN, JAPANESE, show better accuracy performance compared to single models such as *Decision Tree*, Random Forest, Logistic Regression, Support Vector *Machine*, AdaBoost, Gradient Bosting *Machine*, and XGBoost itself. The HMEQ dataset achieved the highest accuracy at 98%, German at 78%, and Japanese at 89%. The *credit scoring* classification model with DNN-XGBoost algorithm using *credit scoring* datasets in this thesis shows better performance than several single models.

Keywords: *Credit Scoring*, Binary Classification, Joint Model, Feature Extraction, *Deep Neural Network*, DNN, Extreme Gradient Boosting, XGBoost.