

ABSTRAK

Kredit merupakan fasilitas keuangan yang memungkinkan individu atau badan usaha dapat melakukan pinjaman yang kemudian akan membayarnya kembali dalam jangka waktu dan bunga pinjaman yang sudah ditentukan. Ada banyak jenis kredit yang disediakan oleh penyedia jasa kredit, salah satunya adalah kredit hipotek (*mortgage loan*) atau yang lebih dikenal dengan Kredit Pemilikan Rumah atau KPR. Pada laporan *Mortgage Finance Forecast* yang dikeluarkan oleh *Mortgage Bankers Association* (MBA) memprediksikan industri pinjaman hipotek untuk memperoleh originasi \$2.5 Triliun setiap tahunnya dalam 3 tahun kedepan atau setidaknya 40% lebih tinggi dibanding dari rata-rata originasi tahunan antara 2010-2019. Pemodelan menggunakan *Machine Learning* diperlukan dalam memprediksi peluang gagal bayar dari nasabah pinjaman hipotek. Dalam penelitian ini digunakan dua jenis metode klasifikasi yaitu *Random Forest* dan *XGBoost* untuk memprediksi peluang gagal bayar nasabah pinjaman hipotek. Hasil penelitian menunjukkan bahwa model *Random Forest* lebih unggul dalam melakukan klasifikasi daripada *XGBoost* dengan dari tingkat akurasi, sensitivitas, dan F1-score. Model *Random Forest* menghasilkan nilai akurasi 78.31%, presisi 68.65%, sensitivitas 75.40%, dan F1-score 71.87%, sedangkan model *XGBoost* menghasilkan nilai akurasi 78.19%, presisi 69.62%, sensitivitas 72.13%, dan F1-score 70.85%.

Kata kunci: *Mortgage Loan, Machine Learning, Random Forest, XGBoost*

ABSTRACT

Credit is a financial instrument that allows individuals or business entities to make loans that will repay them within a predetermined loan period and interest. There are many types of credit provided by credit service providers, one of which is a mortgage loan or a Home Ownership Loan. The Mortgage Finance Forecast report released by the Mortgage Bankers Association (MBA) predicts the mortgage loan industry to earn \$2.5 Trillion in originations annually for the next 3 years or at least 40% more than the average annual origination between 2010-2019. Modeling using Machine Learning is needed in predicting the probability of default from mortgage loan customers. In this study, two types of classification methods were used, namely Random Forest and XGBoost to predict the probability of default on mortgage loan customers. The results showed that the Random Forest model is better at classifying than XGBoost with the level of accuracy, sensitivity, and F1-score. The Random Forest model produces an accuracy value of 78.31%, precision 68.65%, sensitivity 75.40%, and F1-score 71.87%, while the XGBoost model produces an accuracy value of 78.19%, precision 69.62%, sensitivity 72.13%, and F1-score 70.85%.

Keywords: Mortgage Loan, Machine Learning, Random Forest, XGBoost.

