

ABSTRACT

P2P lending is an example of alternative finance and has increased since its establishment in 2005. Consequently, P2P lending is well-known worldwide. P2P gives users, borrowers, and lenders the privilege to choose the manner and the object of their loan. P2P serves transparent data compared to the traditional one. Unfortunately, the growth of the P2P lending industry has exceeded the growth of the law related to fintech. As a result, some concerns are associated with this industry during its development. Since there is no proper regulation, information asymmetry has become a crucial problem in the P2P lending industry. When information between parties is not equally distributed, one of the parties has more information than the others.

This study will try to become an additional literature review for fintech and P2P lending topics. Besides, this study aims to become a practical and theoretical study that proves that artificial intelligence can predict the factor that makes people pay their loans. It will help each player in the P2P lending operations to improve the analysis to reduce default risk and information asymmetry.

The dataset is extracted from LendingClub from 2017 to 2020; LendingClub is known as benchmark data in the P2P lending industry. After learning from the previous literature, the first objective is to improve data quality by doing data pre-processing. In addition, to handle the imbalance issue, SMOTE technique was applied. Several models have been implemented, namely, Logistic Regression, Decision Tree, Random Forest, MLP, KNN and LightGBM, for the data analysis. Evaluation of the model has adopted the confusion matrix series of metrics, and empirical study resulted in LightGBM and Random Forest providing the best performance compared to the other. In terms of evaluation metrics, it predicts with higher accuracy.

Keywords – *peer to peer lending, fintech, artificial intelligence, loan assessment, classification classifiers*