

Title	開発途上国のデータ駆動型信用スコアリングモデルの 解釈可能化 ベトナム海運商業銀行の顧客データの分 析
Author(s)	松永, 智行
Citation	
Issue Date	2021-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/17189
Rights	
Description	Supervisor: Dam Hieu Chi, 先端科学技術研究科, 修 士 (知識科学)

Interpretability of data-driven credit scoring models for developing country: An analysis of customer data of Vietnam Maritime Commercial Joint Stock Bank

1910200 Tomoyuki Matsunaga

Financial services are important in developing countries, and the lack of them is one of the major obstacles for many poor people to improve their living standards. Recently, a number of methods and businesses including FinTech have emerged to address this problem by combining machine learning technology with data such as digital footprints, which had not been used to determine their creditworthiness. However, there is a problem with these methods. That is the anxiety over whether we can really trust the information on creditworthiness (credit score) generated by such new methods. This concern is not limited to the financial sector, and now that AI technology has begun to spread, it is being discussed in various fields as an issue of interpretability in AI (machine learning).

In this study, we examined the need for, and problems of, interpretability in machine learning, which has not been studied in detail in Japan. The insights therefrom were applied to credit scoring model research in the field of finance. In addition, we applied principal component analysis (PCA), clustering, and Association Rule Mining to a data set provided by the Vietnam Maritime Commercial Joint Stock Bank that linked customers' credit scores to their SNS posts. These analyses confirmed some tendencies that existed between the content of SNS posts and credit scores, and contributed, albeit partially, to improving the interpretability of the credit scoring model.

A major question remains: "To what extent should we be able to interpret (explain) the data so that people will trust or accept it? " Our idea of a good explanation is extremely vague and complex due to the problems of truthfulness and confirmation bias, and the point at which an explanation is judged to be good may be affected by people's emotions and interests. Therefore, if the compromise point of "how much explanation is enough" is not appropriate, the cost of credit screening will increase and the premise of the ability to solve financial problems will be weakened. It should be noted that the expansion of FinTech businesses in developing countries was fostered by the looser legal regulations compared to developed countries.