中央研究院 ACADEMIA SINICA 1928
統計科學研究所
INSTITUTE OF STATISTICAL SCIENCE

Seminar

Title：Integrating Explainable AI with Polynomial Analytics to Enhance Credit Scoring Model Compliance

Speaker：Prof. Huei-Wen Teng (鄧惠文 副教授)
(Department of Information Management and Finance, National Yang Ming Chiao Tung University)

Time：10:30 ~ 12:00, Monday, February 26, 2024

Place：Auditorium, B1F, Institute of Statistical Science

Abstract

Adhering to the General Data Protection Regulation (GDPR) (Voigt and Von dem Bussche, 2017) in the European Unions and the Equal Credit Opportunity Act (ECOA) in the United States (Consumer Financial Protection Bureau, 2022), this paper enhances credit scoring models to fulfill both interpretability and accuracy criteria. Although Logistic Regression is known for its interpretability, its accuracy is often limited. We demonstrate that augmenting Logistic Regression with polynomial and interaction features substantially elevates its performance, making it competitive with, or even superior to, the XGBoost algorithm. This improvement, however, raises issues of multicollinearity and overfitting, which we mitigate through a Shapley value-based feature selection method. Experiments on synthetic and open-source datasets corroborate the e↵ectiveness of our enhanced Logistic Regression model. In contrast, XGBoost's performance, in terms of AUC, plateaus with similar feature engineering, underscoring our model's potential as a robust, precise, and interpretable credit scoring tool.

Keywords: Credit Scoring, polynomial and interaction features, Shapley Value, Feature Selection, XGBoost, Logistic Regression
JEL: C51, C52, C53, G21, C38

※ **Tea reception starts at 10：10.**

※ **Online live streaming through Cisco Webex will be available.**