

Abstrakt

Ochrana klientů před podvodnými transakcemi je náročný úkol. Banky se obvykle spoléhají na systémy založené na pravidlech, které vyžadují ruční tvorbu těchto pravidel pro identifikaci podvodu. Tato pravidla musí nastavit zaměstnanci banky, kteří musí sami vyhledávat trendy v podvodných transakcích. Tato práce se zabývá problémem odhalování podvodných karetních transakcí a porovnává několik modelů strojového učení pro detekci podvodů. Tyto modely mohou v datech najít složité vztahy a potenciálně překonat klasické systémy detekce podvodů, Logistická regrese, neuronová síť, random forest a extreme gradient boosting (XGBoost) jsou trénovány na simulovaném souboru dat, který věrně kopíruje vlastnosti skutečných karetních transakcí. Výkonnost modelů se měří podle citlivosti, specificity, preciznosti, AUC a časové náročnosti předpovědi na testovacím souboru dat. XGBoost vykazuje nejvyšší výkonnost mezi testovanými modely. Poté je porovnáván se standardním systémem detekce podvodů používaným v české bance. Bankovní systém dosahuje vyšší specificity, ale XGBoost přesto vykazuje slibné výsledky. Je možné, že některé modely strojového učení by mohly překonat současné systémy detekce podvodů, pokud budou dobře vyladěny.

Klasifikace JEL G21, K42

Klíčová slova strojové učení, karetní podvody, detekce podvodů, nevyvážená data

Název práce Metody strojového učení v detekci podvodných karetních transakcí