

Abstrakt

Hlavním cílem našeho výzkumu bylo vyvinout nový bankrotní model pro českou ekonomiku. Za tím účelem jsme použili logistickou regresi a vzorek 150 000 finančních výkazů pro období 2002—2016. Pracovali jsme celkově s 41 vysvětlujícími proměnnými, z čehož 25 byly finanční poměrové ukazatele a 16 dummy proměnné, k selekci nejlepších prediktorů bylo využito Bayesovské průměrování modelů. Výsledný model byl odhadnut pro 3 predikční horizonty, jeden, dva a tři roky před bankrotem, abychom mohli vyhodnotit vývoj v signifikanci jednotlivých proměnných a přesnost modelů pro různé horizonty. Protože jsme měli významně méně dat pro bankrotující společnosti, použili jsme metody tzv. over-samplingu a under-samplingu pro data, která byla použita k odhadování našich modelů (80% celého vzorku). Tyto metody se ukázaly být velice efektivní, protože zlepšily predikční schopnosti modelů napříč časovými horizonty. Přesnost predikcí jsme měřili pomocí ROC křivek, Sensitivity-Specificity křivek a Precision-Recall křivek. V porovnání s modely odhadnutými na českých datech náš model dopadl velice dobře. Pomocí analýzy ekonomické a statistické signifikance odhadnutých parametrů jsme navíc vybrali nejlepší proměnné pro predikce v krátkodobém a dlouhodobém horizontu, což má přidanou hodnotu pro praktické použití.

Klasifikace

C11, C51, C53, G33, M21

Klíčová slova

Predikce úpadku, Logistická regrese,
Bayesovské průměrování modelů, ROC
křivky