

Student: Mateusz Kuchta

Kierunek studiów: Big Data

Tryb studiów: niestacjonarne

Planowany termin obrony: czerwiec 2023

Proponowany tytuł pracy:

„Wykorzystanie technik analizy predykcyjnej do celów estymacji wiarygodności kredytowej”

Celem pracy jest budowa modelu predykcyjnego/modeli predykcyjnych do wykorzystania w dziedzinie Credit Scoring'u. Rozważone zostanie podejście przy pomocy regresji logistycznej, a także technik machine learning'u, ze szczególnym uwzględnieniem tzw. kontrykcyjnego uczenia maszynowego (ang. adversarial machine learning).

Budowa modelu zostanie poprzedzona szczegółową analizą dotychczasowych dokonań w dziedzinie Credit Scoring'u, zarówno przy wykorzystaniu technik najbardziej popularnych, jak i tych bardziej nietypowych i rzadziej używanych. Następnie zostanie przeprowadzona analiza i dobór odpowiedniego zbioru danych do budowy modeli predykcyjnych. Po dokonaniu eksploracji danych zostanie zrealizowana część praktyczna pracy – realizacja, przetrenowanie, walidacja oraz wykorzystanie modelu predykcyjnego. Wykonana zostanie również analiza porównawcza skuteczności stworzonych modeli.

Do wykonania pracy zostaną wykorzystane technologie takie jak Databricks czy język R, a efekt końcowy zostanie opublikowany przy pomocy aplikacji Shiny.

Praca będzie oparta na literaturze pomocniczej opisującej sens i działanie użytych technik analizy, jak również podręcznik do języka R. Zastosowana zostanie również dokumentacja do współpracy z aplikacją Shiny. Niezbędne będzie również pochylenie się nad literaturą opisującą problem szacowania wiarygodności kredytowej oraz wykorzystanie doświadczeń badaczy w dziedzinie użycia uczenia maszynowego dla dziedziny Credit Scoring'u (np. ["Adversarial Semi-supervised Learning for Corporate Credit Ratings"](#) - Bojing Feng, Wenfang Xue lub ["A Machine Learning Approach for Micro-Credit Scoring"](#) - Apostolos Ampountolas, Titus Nyarko Nde, Paresh Date, Corina Constantinescu).