

Master Universitario in: "ANALISI DATI PER LA BUSINESS INTELLIGENCE"

A.A. 2020/2021

Titolo della tesi: Utilizzo di algoritmi di classificazione in contesto aziendale a supporto della gestione delle pratiche dei crediti deteriorati (Non-Performing Loans)

Autore: Nicolò Mattia Pollara

Abstract

Il presente lavoro ha come obiettivo lo studio e l'implementazione di algoritmi di classificazione in ambito bancario, nello specifico, l'utilizzo dell'algoritmo random forest per la classificazione dei clienti aventi crediti deteriorati. Quanto riportato nell'elaborato è frutto del lavoro svolto durante il tirocinio curriculare dell'autore presso l'ufficio di Machine Learning & Artificial Intelligence di Intesa Sanpaolo. Durante questo periodo l'autore ha realizzato un prototipo di modello decisionale che sia di supporto per la gestione delle pratiche dei crediti deteriorati (noti anche come Non-Performing Loans): viene descritto nell'elaborato il processo di approvvigionamento, pulizia e trattazione dei dati relativi ai clienti al fine di realizzare le feature del modello predittivo. A partire dai dati relativi alle pratiche dei clienti, il modello costruito è in grado di effettuare una classificazione delle nuove pratiche in due categorie: clienti da seguire con priorità avendo maggior probabilità di rientro e clienti con minor probabilità di rientro della situazione di passività. Il corpo centrale della tesi riporta i concetti base che costituiscono le fondamenta teorica degli alberi decisionali, indicandone i criteri di creazione di tali modelli oltre al loro funzionamento, con il fine di presentare successivamente una loro estensione: l'algoritmo random forest. Di questo, sono presentate le caratteristiche ed il funzionamento. La sezione finale dell'elaborato riporta i risultati dell'applicazione di diversi algoritmi di classificazione al problema precedentemente descritto, mettendoli a confronto sulle metriche di accuratezza, recupero e precisione. Dal confronto emerge la superiorità nelle performance dell'applicazione dell'algoritmo random forest e di questo vengono trattate ulteriori considerazioni con il fine di utilizzare il prototipo implementante come supporto nella gestione delle pratiche. I benefici introdotti dal prototipo e le sue future evoluzioni sono rendicontati nel capitolo finale dell'elaborato.