

Následující diplomová práce zkoumá vliv charakteristik dat z hlediska jejich rozdělení na výkonost klasifikačních metod. Používáme metodu Gassovských kopulí a simulaci 1000 umělých datasetů. Trénujeme klasifikační metody GLM, GBM, DRF, XRT pomocí platformy pro strojové učení H2O.ai, přistupujeme k ní však pomocí R. Na těchto datasetech testujeme výkonost jednotlivých klasifikačních metod a prezentujeme vyzorované souvislosti. Dále používáme reálný dataset o Australských půjčkách a předpovíáme, která by měla být ta nejvýkonější. Predikovaná výkonost jednotlivých metod je počítána pomocí penalizování za rozdílnost Australského datasetu od datasetů, kde každá z metod fungovala komparativně lépe. Tento způsob předpovědi se ukázal jako nefungující.