Introduction into Big Data analytics

Assignment 5

Janusz Szwabiński

Task 1 Download the Bank Marketing dataset from Kaggle (`https://www.kaggle.com/rouseguy/bankbalanced/data`). Using PySpark, build a binary classifier that predicts whether a client will subscribe to a term deposit. Compare different classification algorithms with each other.