

Comparison between XGBoost, LightGBM and CatBoost Using a Home Credit Dataset

Authors : Essam Al Daoud

Abstract : Gradient boosting methods have been proven to be a very important strategy. Many successful machine learning solutions were developed using the XGBoost and its derivatives. The aim of this study is to investigate and compare the efficiency of three gradient methods. Home credit dataset is used in this work which contains 219 features and 356251 records. However, new features are generated and several techniques are used to rank and select the best features. The implementation indicates that the LightGBM is faster and more accurate than CatBoost and XGBoost using variant number of features and records.

Keywords : gradient boosting, XGBoost, LightGBM, CatBoost, home credit

Conference Title : ICAIDM 2019 : International Conference on Artificial Intelligence and Data Mining

Conference Location : Bangkok, Thailand

Conference Dates : February 04-05, 2019