

БВИ НЕ БУДЕТ

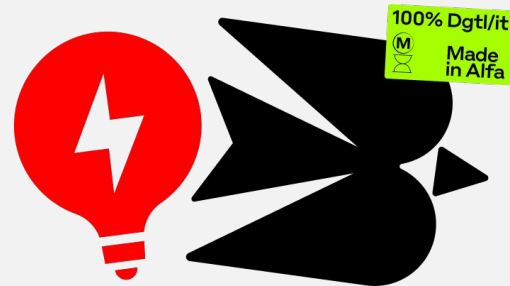
Обучение моделей для прогнозирования кредитоспособности



БВИ не будет

Подзаголовок

Влад Бухарин
Игорь Новолодский
Максим Гузюк
Кирилл Пучков



Модели

Random Forest

Random Forest - алгоритм, использующий деревья решений для прогнозирования результатов, вместе с Grid Search, возможно комбинировать несколько моделей и еще больше улучшить точность.

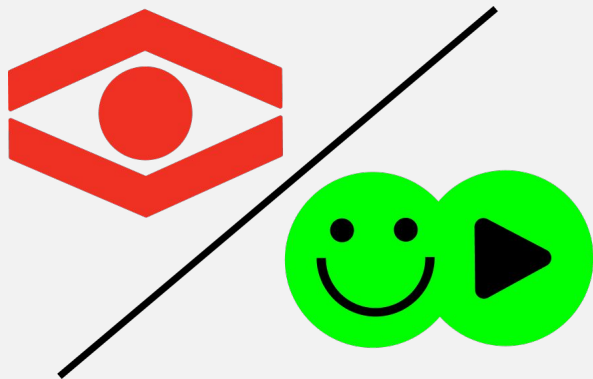
```
clf = RandomForestClassifier(random_state=42)
grid = {
    "n_estimators": [100, 200, 300],
    "max_depth": [5, 10, 15],
    "min_samples_split": [2, 5, 10]
}
model2 = GridSearchCV(clf, param_grid=grid, cv=5)
model2.fit(x_train, y_train['flag'])
```

Catboost

CatBoost - это мощный алгоритм, который хорошо справляется с обработкой как числовых, так и категориальных данных. В отличие от Random Forest, деревья решений применяются последовательно, ошибки предыдущих моделей корректируются последующими.

```
model1 = cb.CatBoostRegressor(loss_function="Poisson")
grid = {'iterations': [192, 193],
        'learning_rate': [0.1],
        'depth': [7, 9],
        'l2_leaf_reg': [3, 4]}
model1.grid_search(grid, dataset)
```

Признаки



Статистические функции

- Min, Max
- Mean, Median
- Count, Num Unique Values

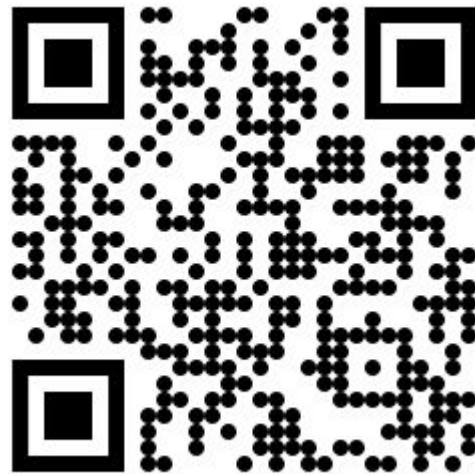
Выделение уникальных категорий

- City, Country
- MCC, MCC Category, eCommerce Flag
- Operation Type, Payment System, Income Flag
- И другие



GitHub

Все файлы проекта можно найти тут



github.com/HghaVlad/ML_alfacampus