CLASSIFICATION & REGRESSION MODEL

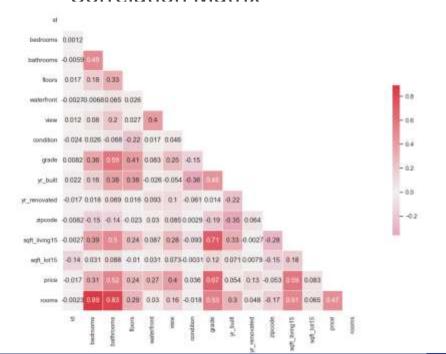
Javier Miguélez, Juan Antonio Pérez y Diego de Santos

Regression Model: Data Cleaning



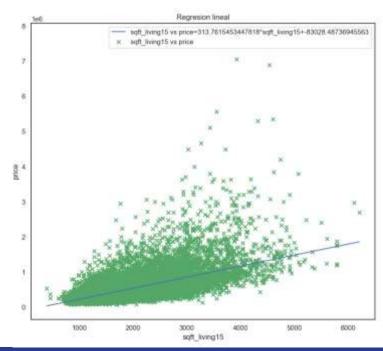
Regression Model: Análisis

Correlation Matrix

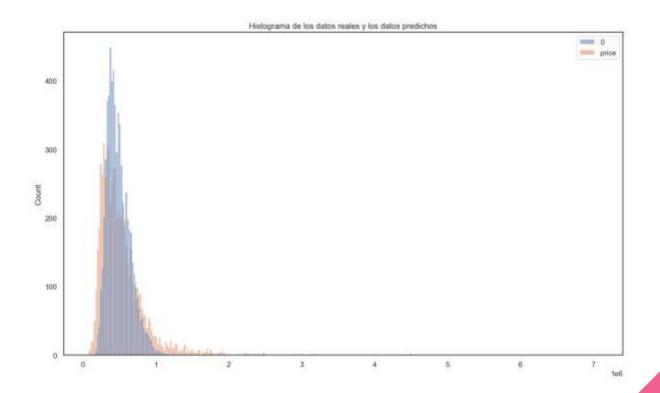


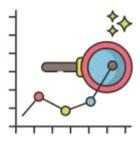


Linear Regression



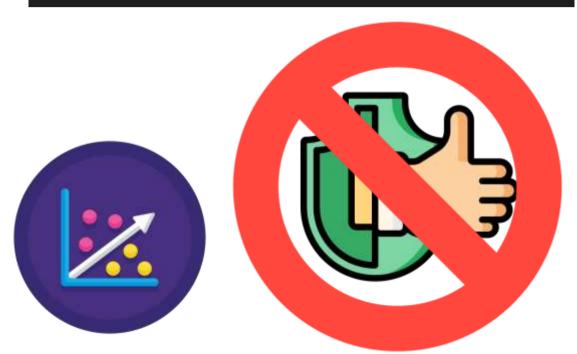
Regression Model: Predicción





Regression Model: Conclusión y Resultados

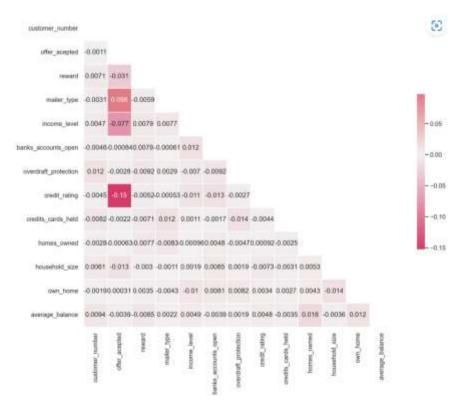
El R2 es de: 0.27841310195830504

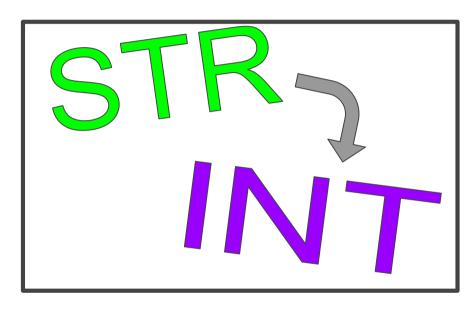






Classification Model: Data cleaning



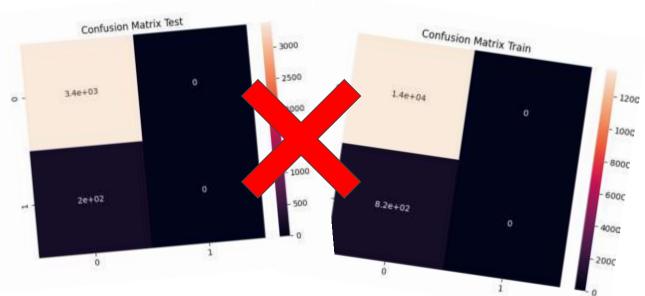


Classification Models: Primeros Modelos

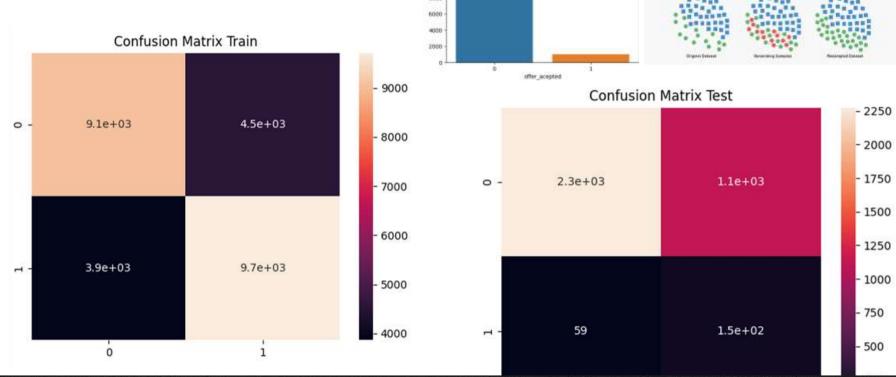


Test









14000

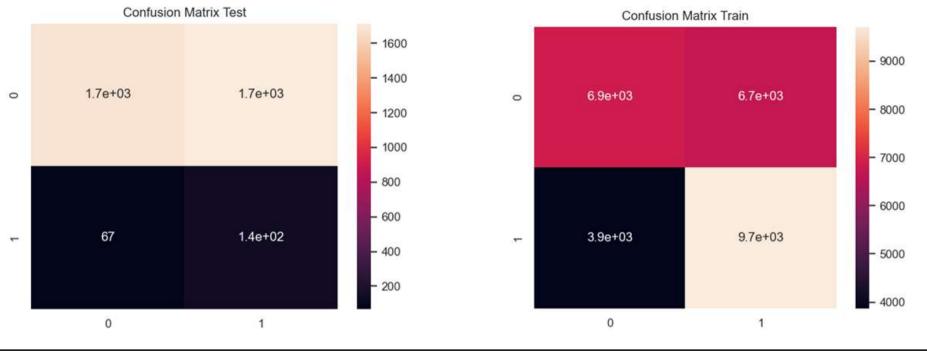
12000

SMOTE

HANDLE IMBALANCED DATASET

Synthetic Minority Oversampling Technique

{'le_train_sm_score': 0.6836990134000883, 'le_test_sm_score': 0.6872222222222222, 'le_train_precision': 0.6782652186338954, 'le_test_precision': 0.11973575557390587, 'le_train_recall': 0.698939773229274, 'le_test_recall': 0.7073170731707317, 'le_f1_train_sm': 0.68844731307564, 'le_f1_test_sm': 0.20480225988700568}



{'lr_train_score': 0.6072007068178471, 'lr_test_score': 0.514444444444445, 'lr_train_precision': 0.5888129803586678, 'lr_test_pre cision': 0.07679648930334614, 'lr_train_recall': 0.7107200706817847, 'lr_test_recall': 0.6829268292682927, 'lr_f1_train': 0.644048 5721910861, 'lr_f1_test': 0.13806706114398423}

Classification Model: ¿Otro enfoque?

Streamlit



