Bitcoin Address Detection

Asrar Abdulrhman
Mozah Alkhaldi

INTRODUCTION

Bitcoin is a cryptocurrency based on peer-to-peer technology that involves no central authority like a bank. For all intents and purposes, transactions done over bitcoin cannot be traced.

Problem statement & Goal

Given a bitcoin address along with some meta-data pertaining to that address,

we are challenged to predict if that address has been used to receive ransoms in the past.

Data Set

The data to be tested in this project is from UCI Machine Learning Repository archive-beta.ics.uci.edu

Original data：(2916697, 10)

Tools & Data Set

Tools

pandas

NumPy

matplotlib

seaborn

statsmodels

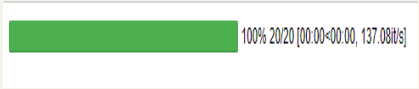
scikit learn

Pickle

Pre-processing Tools:

boxcox

tqdm



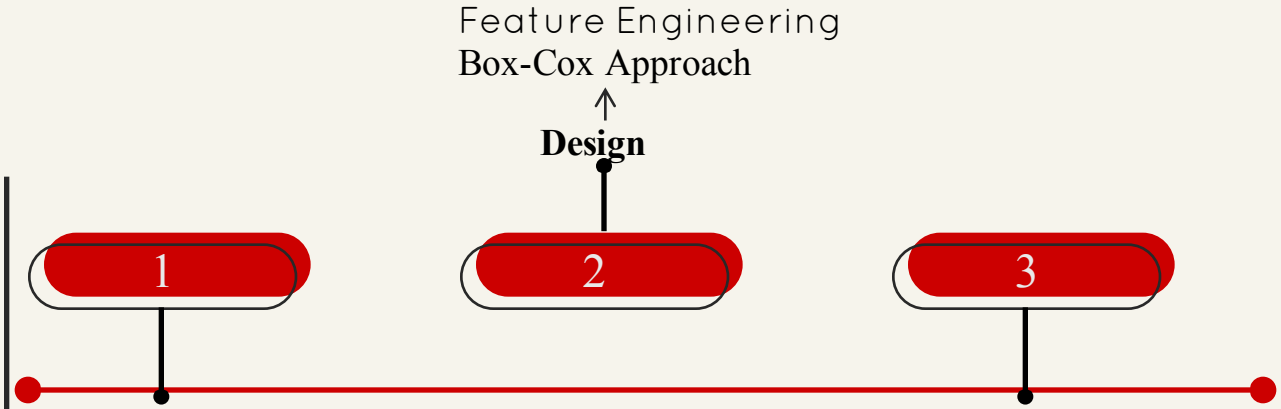
datetime

model_selection

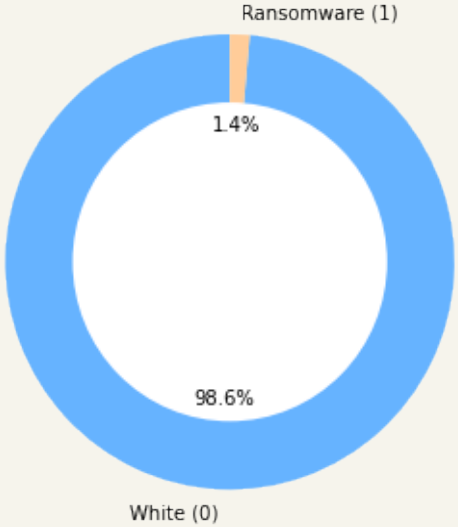
Voting Approach :

StackingClassifier

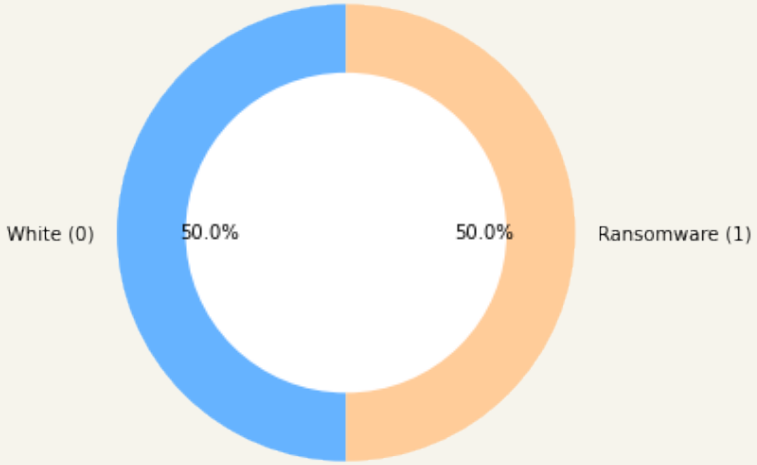
Project Workflow



Handling Ensembling issue

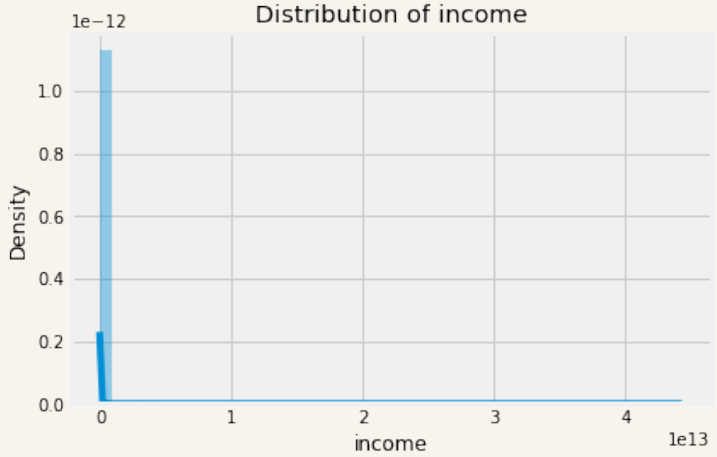


Before Undersampling

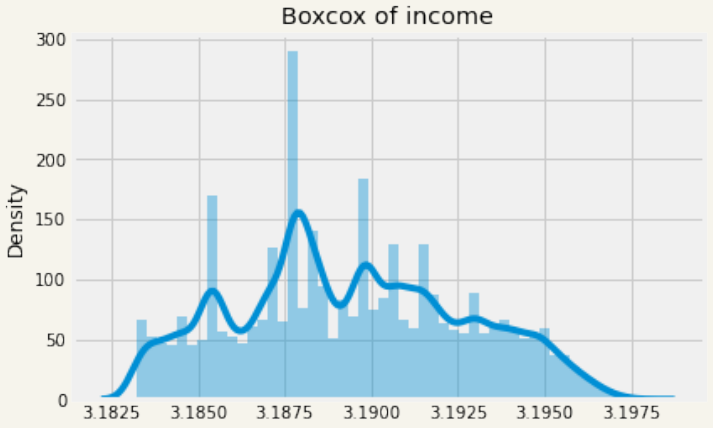


After Undersampling

DESIGN



Before using Box-Cox



After using Box-Cox

MODEL SCORES

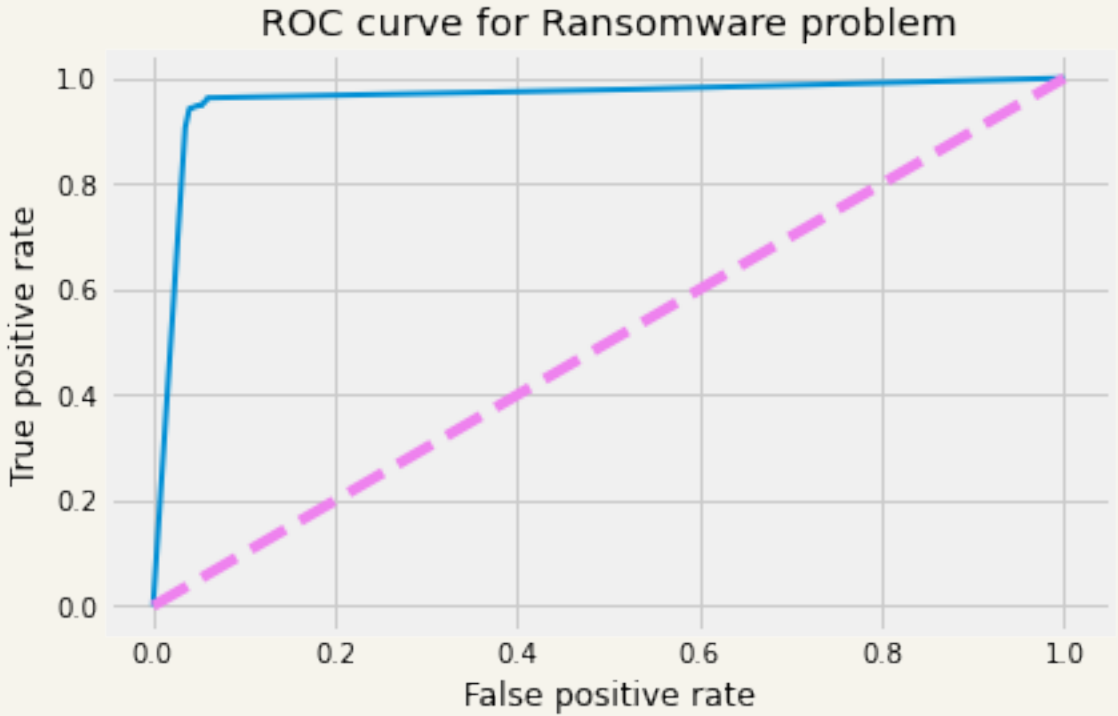
Classifier	Accuracy	Recall	Precision	F score	F(beta=2)
LogisticRegression	0.74658	0.6696	0.7912	0.7253696	0.69098295
KNeighbors	0.891283	0.9384	0.8575	0.8961301	0.92102143
RandomForest	0.955088	0.9554	0.9547	0.9550833	0.95529093
ExtraTrees	0.9476638	0.9459	0.9496	0.9477427	0.948875156
BernoulliNB()	0.756549	0.6361	0.8377	0.7231033	0.668240190
GaussianNB()	0.600507	0.9768	0.5572	0.7096349	0.848957567

Voting Approach: Stack Classifier

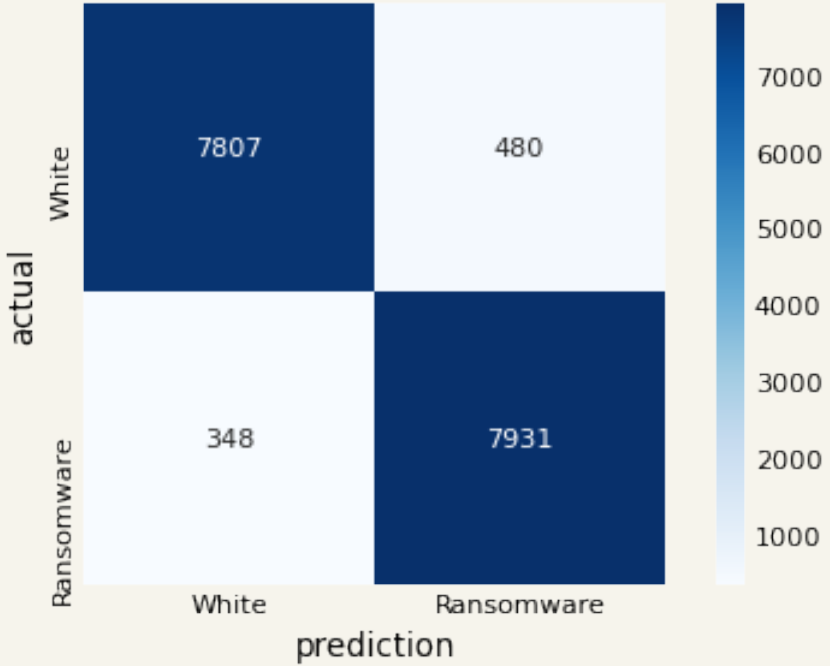
Classifier	Accuracy	Recall	Precision	F score	F(beta=2)
Stack	0.95001810	0.9768	0.5572	0.709634	0.84895756

meta_classifier=RandomForestClassifier()

ROC curve On Stack Classifier



Confusion Matrix on Stack classifier



Thank you

For listening