MONEY LAUNDERING DETECTION IN CRYPTOCURRENCY NETWORKS

by

Elif Emine Erdem

B.S., Industrial Engineering, TOBB University of Economics and Technology, 2017

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Industrial Engineering

Boğaziçi University

2022

ACKNOWLEDGEMENTS

I am incredibly grateful to my professor Tınaz Ekim Aşıcı for her helpful patience and feedbacks. Additionally, without the extraordinary assistance of Tolga Kurt, Utku Görkem Ketenci, and Hikmet Mazmanoğlu from H3M, this work would not have been accomplished. Moreover, without the knowledge and experience that my defense committee Yaşar Safkan, Mustafa Gökçe Baydoğan so kindly provided, I would not have been able to go on this journey.

I would like to thank my father Şinasi for giving me a happy and loving childhood. My father will always be an inspiration for me in life because of his courage, determination, and love. I would also like to thank my mother, Fevziye, who showed me what a woman could do and achieve alone. Not saying thank you to my siblings would be foolish, Batuhan and Ayşe, whose beliefs and wisdom always motivated me to become a better person.

ABSTRACT

MONEY LAUNDERING DETECTION IN CRYPTOCURRENCY NETWORKS

This study aims to develop scalable methods to detect suspicious wallets using historical transaction data in cryptocurrency networks such as Ethereum and Bitcoin. Different transaction networks are generated for each wallet data set using the illicit wallets dispersed around the internet. Egonet-dependent and independent features are used with a range of machine learning techniques, including logistic regression (LR), random forest (RF), and XGBoost (XGB), to predict illicit wallets. Firstly, we analyze performance of models to detect suspicious wallets in the two datasets that include suspicious bitcoin mixer services wallets such as Bitcoinfog and Helix. The area under the ROC curve value (AUC) is over 99% for XGB models. We observe that models perform better on Helix wallets than BitcoinFog wallets in terms of precision, recall, f1 score, and AUC. Secondly, we notice that egonet dependent features do not significantly improve the models' performances. Hence, best-performing models have only egonet independent features. Thirdly, on Bitcoin datasets that do not use any mixer services, we obtain over 99% AUC. Although the performance of the models is similar in these three datasets, dominant features in terms of feature importance measure are different between the datasets including wallets using mixer services (Helix, Bitcoinfog) and the other (Bitcoin). Lastly, utilizing the same feature set as we do on Bitcoin, Bitcoinfog and Helix datasets, we train the same machine learning models on the Ethereum dataset and obtain 96% AUC. We repeated the tests with varying degrees of class imbalance to simulate real-life situations. We observe a decline in AUC up to 0.10 together with the increasing severity of the class imbalance.

ÖZET

KRİPTOPARA AĞLARINDA KARA PARA AKLAMA FAALİYETLERİNİN TESPİTİ

Bu çalışma, Ethereum ve Bitcoin gibi kripto para ağlarında geçmiş işlem verilerini kullanarak şüpheli cüzdanları tespit etmek için ölçeklenebilir yöntemler geliştirmeyi amaçlamaktadır. İnternete yayılmış yasa dışı cüzdanlar kullanılarak her bir cüzdan veri seti için farklı işlem ağları oluşturulur. Egonet'e bağlı ve bağımsız özellikler, yasadışı cüzdanları tahmin etmek için lojistik regresyon (LR), rastgele orman (RF) ve XGBoost (XGB) dahil olmak üzere bir dizi makine öğrenme tekniğiyle birlikte kullanılır. İlk olarak, Bitcoinfog ve Helix gibi şüpheli bitcoin mikser hizmetleri cüzdanlarını içeren iki veri setindeki şüpheli cüzdanları tespit etmek için modellerin performansını analiz ediyoruz. ROC eğrisi değerinin (AUC) altındaki alan, XGB modelleri için %99'un üzerindedir. Modellerin Helix cüzdanlarında, kesinlik, duyarlılık, f1 puanı ve AUC açısından BitcoinFog cüzdanlarından daha iyi performans gösterdiğini gözlemliyoruz. İkinci olarak, egonet bağımlı özelliklerin modellerin performanslarını önemli ölçüde iyileştirmediğini fark ettik. Bu nedenle, en iyi performans gösteren modeller yalnızca egonetten bağımsız özelliklere sahiptir. Üçüncüsü, herhangi bir mikser hizmeti kullanmayan Bitcoin veri setlerinde %99 AUC elde ediyoruz. Modellerin performansı bu üç veri setinde benzer olsa da, mikser hizmetlerini kullanan cüzdanlar (Helix, Bitcoinfog) ve diğer (Bitcoin) içeren veri setleri arasında özellik önem ölçüsü açısından baskın özellikler farklıdır. Son olarak, Bitcoin, Bitcoinfog ve Helix veri setlerinde kullandığımız aynı özellik setini kullanarak, aynı makine öğrenme modellerini Ethereum veri setinde eğitiyor ve %96 AUC elde ediyoruz. Gerçek yaşam durumlarını simüle etmek için testleri değişen derecelerde sınıf dengesizliği ile tekrarladık. Sınıf dengesizliğinin artan şiddeti ile birlikte AUC'de 0,10'a kadar bir düşüş gözlemliyoruz.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Cryptocurrency is defined as a digital currency in which transactions are verified and records are kept by a decentralized system using cryptography rather than a central authority [1]. A digital currency using cryptography was first proposed by Chaum in 1983 [2]. Satoshi Nakamoto proposed Bitcoin in 2008, which is the first decentralized virtual currency [3]. Today, there are more than 10,000 cryptocurrencies like Bitcoin, Ethereum, Ripple, Litecoin, NEO, and so on. Among them, Bitcoin and Ethereum have the highest trading volumes.

It is considered that 2 billion dollars are laundered throughout the world each year. However, only 2% of it can be detected by regulators and law enforcement [4]. After emerging cryptocurrency, criminals have begun to exploit it for money laundering activities. Liberty Reserve is one example of them [5]. It laundered approximately 6 billion dollars by utilizing the anonymous and untraceable characteristics of digital currency. Besides, a considerable number of drug dealers in Costa Rica used Liberty Reserve for illegal activities for many years.

Financial institutions such as banks, brokerage firms, and investment dealers are responsible for the security of the transactions. Buyers and sellers give extensive authority to financial institutions to protect themselves from malicious activities. Thus, one cannot make transaction without their permission. If a financial institution witnesses a malicious or illicit activity, it can reverse the transaction to prevent it. On the other hand, cryptocurrencies do not have a third-party regulator due to their decentralized nature. Individuals can have more chances to hide in the network and make a transaction without being tracked.

The decentralized nature of cryptocurrencies offers people full control over their privacy. However, money launderers make use of this golden opportunity to hide their identities. Yet, illicit activities in cryptocurrency can still be detected since

transaction networks are publicly available via blockchain since a certain degree of user anonymity is available in cryptocurrency. First, the transaction flow may be clearly seen by anybody on the blockchain, which increases the likelihood of a deanonymization assault by grouping linked addresses into a single wallet [6–10]. Second, Know Your Customer (KYC) policies are commonplace across exchanges. Lastly, it is possible to deanonymize individuals by leveraging publicly accessible information in blockchain networks [9, 11, 12]. Besides, Bitcoin transactions may include simple connections that are easy to track. In order to increase anonymity, the Bitcoin ecosystem has developed mixer services. Multiple Bitcoin transactions are mixed together in a mixer so that the connections between the transactions are concealed. Mixing services mix the funds provided by parties and distribute them to different wallets to hide illicit activities by severing the links between wallets. When there are several input-output wallets, the mixer combines them so that the link of input and output wallets are lost. As a result of this, new links between the input-output wallets are generated [9]. In Section 2.1.1.1, we explained Bitcoin mixer services detailed. The number of wallets using mixing services has increased recently. According to Crystal Blockchain End of Year Report 2020, in 2020, mixers received a total of $1.4 billion in bitcoin, compared to $0.9 billion in 2019. This means that throughout the course of a year, the amount received by the mixers grew by 0.5 billion dollars [13].

If a mixing service uses a central mixing server to conduct the mixing, it is referred to as a centralized mixing service. However, there is no assurance that service providers would transmit mixed currencies to user-specified addresses. Also, they can keep track of the initial connection between user inputs and outcomes [14]. Therefore, applying for a mixer service includes risks. The first mixing services platform, BitcoinFog, has dedicated servers that are manually set up. The service is only accessible via Tor. A qualified workforce keeps up the mixing services to prevent the huge mixing pool from hazardous assaults [15]. Nevertheless, money launderers apply for mixers to cover their tracks.

Recently, studies on the anonymity of Bitcoin transactions using Blockchain technology have been expanded [16]. In the Bitcoin network, a single transaction contains numerous money transfers that are made between different wallets. As a result of this, the Ethereum network has been started to highlight in recent studies [17, 18]. Despite the fact that Ethereum blocks include many transactions, each transaction only comprises a receiver and a sender. As a consequence, determining which wallets performed the transaction is explicit in the Ethereum network. All these recent developments make it crucial to have more powerful and adaptive methods to detect suspicious cryptocurrency transactions. The study has a variety of goals:

- ➢ Developing a model that predicts suspicious bitcoin mixer wallets in accordance with current bitcoin mixer activity. This might be used to discover unknown illicit mixer wallets on the bitcoin network.
- ➢ Comparison of the effects of independent and dependent egonet features on models.
- ➢ Comparing models that detect illegal wallets using Bitcoin mixer services such as BitcoinFog and Helix.
- ➢ Developing a model to detect suspicious Bitcoin wallets and evaluating the results using the previous model.
- ➢ Developing a model that detects illicit Ethereum wallets, detecting similar wallets, comparing the outcomes of this model to the two previous ones, and identifying significant features.
- ➢ Experimenting with various illicit/licit wallet ratios and evaluating outcomes across all models.

In the following sections, detailed information on cryptocurrencies, including bitcoin, Ethereum, and bitcoin mixer services, is provided in Section 2. This section also gives brief information on the basic areas where machine learning and artificial intelligence are used in crypto money networks. Additionally, studies on money laundering detection in cryptocurrency research are detailed in Section 2. The methods for gathering transaction networks of licit and illicit wallets are explored in Section

3 while feature extraction is covered in Section 4. Subsequently, the implementation of our experimental design is covered beginning in Section 5.1, and the outcomes are presented in Section 5.2. In Section 5.2, we begin with our evaluation metrics in results before reporting the findings of each experiment, which are then further addressed in Section 5.2.3. Finally, in Section 6, we summarize our key results, the major problems we encountered, and the prospects for future study in money laundering detection in crytocurrency networks.

2. RELATED WORK

In this section, the most recent research on anti-money laundering practices in Bitcoin networks is given. Along with information on money laundering operations, there is also information on cryptocurrencies like Bitcoin, Ethereum, and Bitcoin mixer services. This section also includes references to earlier theoretical and applied works similar to our study.

2.1. Cryptocurrency

A decentralized system utilizing cryptography, as opposed to a centralized one, is used to verify transactions and maintain records in a digital currency known as cryptocurrency [1]. The concept for a digital money based on cryptography was first proposed by Chaum in 1983 [2]. After that, he designed the eCash digital money system. The fact that eCash was centralized through banks is one of the most key differentiators between cryptocurrencies and the digital money system envisioned by Chaum [19]. While leveraging cryptography technology, cryptocurrency is generated with the use of blockchain technology.

A blockchain is a sort of Digital Ledger Technology (DLT) that is made up of an expanding list of data, known as blocks, that are safely connected to one another using encryption [20]. Satoshi Nakamoto, the person who created Bitcoin, is credited with inventing the blockchain. According to Nakamoto, Bitcoin needed a reliable third party in order to work as a cash without applying for traditional financial systems such as banks. Therefore, Bitcoin must avoid various errors, including transferring money to the wrong account and preventing double spending by the same user. Blockchain take the position of the reliable third party. Blockchain, a database that records every bitcoin transaction, gives evidence of who is in charge of what at any one time [21]. In cryptocurrencies, different amounts of validated transactions constitute each block. Each cryptocurrency system has a predetermined and capped maximum block size [19].

In cryptocurrency's world, transactions must be verified before being added to the public ledgers, which is the responsibility of miners. Since a user could attempt to verify an invalid transaction, each new transaction is examined by the miner to determine if it includes any questionable activities such as using same currency more than once [19]. The two cryptocurrencies with the highest trading volume today are Bitcoin and Ethereum.

2.1.1. Bitcoin

As we mentioned above, the first decentralized virtual money, Bitcoin, was suggested by Satoshi Nakamoto in 2008 [3]. There was initially merely 50 Bitcoin in circulation. The whole amount was almost equal to one US dollar [22]. As the popularity of the use of Bitcoin increases, the value of Bitcoin increases, too. The price at the time this thesis was written was 19,740.61 in US dollars [23]. One of the key aspects of Bitcoin, like other types of currency is scarcity [24]. As a result, Bitcoin's value tends to rise over time.

According to Fauzi et al., there are three advantages for using Bitcoin: security, transaction cost and high return. First, since Bitcoin utilize blockchain technology, it removes the need for intermediary organizations. Besides, confidential information can be safeguarded with this way. Therefore, Bitcoin is more secure than traditional money. Second, in terms of transaction cost, the cost of Bitcoin transaction is lower compared to other normal currencies. Additionally, through the Internet, people can conduct transactions without time constraints and pay a low fee. Third, having a high return increases the popularity of Bitcoin since some investors may benefit by purchasing it at a discount and then selling it at a high [22].

2.1.1.1. Bitcoin Mixer Services. As it is mentioned in the introduction, although Bitcoin offers its users transaction privacy, there are specific circumstances in which this anonymity may be violated. This is so because the transaction flow can be seen by anybody using the blockchain, and most exchange services adhere to their KYC poli-

cies. We cannot conclusively claim that a user of bitcoin mixer services has performed a suspicious transaction. Hence, users may prefer mixer services to hide the flow of bitcoin transactions. However, a number of studies have revealed that mixing services are regularly misused for illicit activities [9]. As previously said, mixer services make it harder to discover the link between input-output pairs in the Bitcoin network. Figure 2.1 depicts a basic bitcoin transaction. The relationship between addresses is explicit in this transaction. A, B, and C send 20, 15, and 5 bitcoins to X, Y, and Z respectively. When the mixer algorithm is implemented, user-A would normally transmit 20 bitcoins to user-X, but user-X would receive the 20 bitcoins in batches from B and user-C rather than straight from user-A to conceal the relationship (see in Figure 2.2).

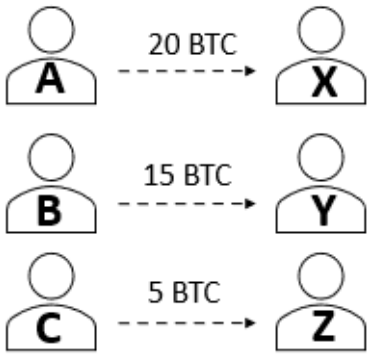

Figure 2.1. A Bitcoin network.

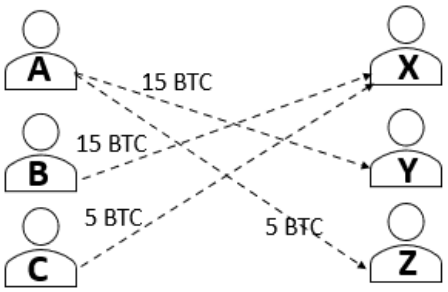

Figure 2.2. The Bitcoin network after mixer service.

In this thesis, we use BitcoinFog and Helix mixer services wallets. BitcoinFog and Helix is centralised mixer services. In December 2013, a portion of the 96,000 BTC from the Sheep Marketplace fraud was cleaned up using BitcoinFog. Moreover, 7,170 bitcoins total were taken from the Chinese exchange Bter.com in February 2015. The theft was linked to some mixer services such as Bitcoinfog. Federal agents from the United States arrested Bitcoinfog's founder in April 2021. It was claimed that during the course of its ten years of existence, Bitcoinfog allegedly laundered over 1.2 million Bitcoin worth about 335 million dollars [25]. The US Federal Bureau of Investigation (FBI) claims that Helix transferred over 350,000 bitcoins on behalf of clients. Darknet markets accounted for the majority of trade volume [26].

2.1.2. Ethereum

The creator of Ethereum is Vitalik Buterin in 2014 [27]. The main difference between Bitcoin and Ethereum is blockchain technology that they use. Vujičić's study reveals how the two cryptocurrencies' blockchain systems differ from one another [28].

2.2. Machine Learning and Artifical Intelligence Implementions in Cryptocurrency Networks

In this part, we looked into the applications of machine learning in cryptocurrency world.

2.2.1. Price Prediction

One of the most important issues for digital currencies which utilize machine learning models is price prediction. Chen has presented a Deep Neural Network (DNN) technique for predicting the price of Bitcoin [29]. The suggested technique produced valuable results in terms of accuracy (53.4%) and correct prediction (MSE 1.02). Another research conducted by Alessandretti et al. is also about cryptocurrency price prediction [30]. All currencies with a history of more than 50 days from their first

appearance and a volume of more than 100000 dollars are included in this study. They focused on three supervised approaches for short-term price forecasting. They are XGBoost with one single regression model, long short-term memory networks (LSTM), and XGBoost with multiple regression models. Moreover, Akyildirim et al. presented a study about the prediction of returns [31]. They concluded that the trend of returns in cryptocurrency markets can be forecasted over daily or minute timeframes. The average accuracy result of all models they used is around 55–60%. They have twelve cryptocurrencies that dominate the marketplace 79.8% of the overall cryptocurrency market to examine. These are Bitcoin, Cash (BCH), Bitcoin (BTC), Dash (DSH), EOS (EOS), Ethereum Classic (ETC), Ethereum (ETH), Iota (IoT), Litecoin (LTC), OmiseGO (OMG), Monero (XMR), Ripple (XRP), and Zcash (ZEC).

2.2.2. Trading

Koker et al. conducted a model using the reinforcement model to obtain a profitable buy-and-hold approach [32]. They tested the model on five main cryptocurrencies in circulation. At the end of the tests, they demonstrate how their active trading approach may improve investors' risk-adjusted returns over time and if the model is properly trained even lower portfolio downside risk.

According to Attanasio et al., multi-coin trading greatly enhances total profits compared to trading only Bitcoin [33]. They also emphasized that there are limited studies about trading algorithms using other digital money except for Bitcoin. They utilize machine learning model to forecast the price of different cryptocurrencies the next day. They use price prediction models intead of classification techniques. Besides, they studied an eight-year dataset for their work.

Sattarov et al. use a deep reinforcement learning model to increase profits from bitcoin transactions [34]. According to the Bitcoin trial, the investor made 14.4% net earnings in just one month. The results of testing on Litecoin and Ethereum were likewise successful, with profits of 74% and 41%, respectively.

2.2.3. Money Laundering Detection

Today, cryptocurrencies drive the attention of many researchers to detect illicit activities in decentralized cryptocurrency systems. Aziz, Baluch, Patel, and Ganie used the Light Gradient Boosting Machine (LGBM) technique to detect Ethereum illegal transactions and they evaluate the performance of several machine learning models in their research. They use the data set which is compiled using Ethereum Classic available on the Kaggle website [35].

There are several studies proposing methods to find illegal activities in cryptocurrency networks. Suranga et al. collected data over three years. They included many wallets for regular and Bitcoin mixing services [36]. All transactions connected to these wallets were gathered. They labeled the Bitcoin mixing services data based on existing literature, news articles, and wallet tags from trusted online resources. In the feature engineering phase, they used in-degree/out-degree ratio, sum, mean, standard deviation of output values, several weakly connected components, and the size of the subgraph a transaction belongs to. As a feature importance method, they preferred the Adaptive Boost (Adaboost) techniques. They tried Neighbourhood, Deepwalk, Node2vec, "OR" ensemble, "AND" ensemble ML techniques to detect suspicious transactions. After some experimental runs, they obtained more accurate results from the Node2vec.

Vassallo et al. worked with Elliptic Data Public which is found in Kaggle, and Ethereum Illicit wallets data flagged by the Ethereum community [18]. They identified three primary areas which require further exploration: which ensemble is most effective in detecting illicit activities on cryptocurrency networks, data sampling techniques, and handling of concept drift in cryptocurrency transactional data. They proposed an innovative adaptation of XGBoost, coined as Adaptive Stacked Extreme Gradient Boosting (ASXGB) to improve the handling of concept drift. To detect money laundering at a transactional level, they used XGBoost due to its efficiency, scalability, and ability to reduce training time by utilizing the GPU. Also, they provided a categorization of machine learning techniques used in the literature.

Li et al. first gathered a sizable dataset of illegal wallets in order to identify the criminal wallets in the Bitcoin network [37]. The illegal wallets primarily originate from a few particular websites, and public forums. Then, they analyzed the features using a variety of machine learning methods (RF, SVM, XGB, ANN), which shows that the suggested features are robust and discriminative. The study also covers the issue of class imbalance.

Lee et al. extracted 80 attributes relating to bitcoin wallets [38]. As classification models, they utilized the random forest and ANN algorithms. The accuracy of the random forest algorithm was much greater than that of the ANN algorithm. In addition to that, they used temporal features by integrating LSTM into the auto-encoder.

Chen et al. studied hackers who try to steal Bitcoin wallet keys to transfer Bitcoin from compromised users [39]. In this work, supervised learning techniques are utilized to find and issue warnings about Bitcoin theft incidents in order to wallet the security concern of Bitcoin theft. They used many algorithms such as k-nearest neighbor (KNN), support vector machine (SVM), random forest (RF), adaptive boosting (AdaBoost), multi-layer perceptron (MLP) techniques, and multiple unsupervised methods. At the end of the study, the RF algorithm was identified as the most effective of these algorithms,

Another study which was proposed by Camino et al. was about two different case studies: money laundering on bank transactions, and Ripple which is another cryptocurrency [40]. They examined three different methods: Isolation Forest (IF), One-Class SVM (OCSVM), and Gaussian Mixture Models (GMM). After getting the anomaly score of each method, they calculated their average to find the suspect list.

There are also other studies which used ML techniques to identify suspicious activites in cryptocurrency network [36, 41–46].

A different research project plans to detect anomalies within graph-based data. Before the model conduction, the authors clustered Bitcoin transactions into predefined particular segments. Segments are exchange, gambling, hosted wallet, merchant services, mining pool, mixing, ransomware, scam, tor market, or other [47]. The authors worked with Chainalysis, a business that specializes in blockchain analysis. After gathering data from Chainalysis, they tested several machine learning methods such as random forests, extremely randomized trees, and adaptive and gradient boosting. At the end of the study, they obtained an average cross-validation accuracy of 80.42 percent by using the gradient boosting method with default parameters [47]. Jung et al. proposed an approach to detect Ponzi schemes in Ethereum transactions [48]. They gathered 172 Ponzi scheme wallets which are found at [1]. 0-day features and behavior-based features are the two types of features that are used in the study to improve the model. The bytecode and the size of the smart contract are examples of the 0-day features whereas paying the early investors from the payments later investors make is an example of behavior-based features. They tested three classification models such as Random Forest, J48, and Stochastic Gradient Descent.

There are statistical challenges such as different data distributions and data amounts among participants. Therefore, the data is not homogenous. At the conclusion of the training procedure, the model accuracy could be poor since the distribution of the data has a significant influence on the model's accuracy [49]. It can be sought to examine methods used in financial transaction networks, such as clustering sub-graphs [50], finding sub-graphs similar to suspicious sub-graphs [51], and identifying suspicious sub-graphs in the networks [52].

Although there are many methods in the literature, there are also state-of-the-art money laundering methods in fast-developing crypto networks. Thus, methods that detect money laundering activities previously found in the literature may be insufficient since Bitmixing services such as Bitcoinfog develop increasingly advanced methods to complicate their transactions. Besides that, in the majority of research in the literature, suspicious wallets that are found on the internet are often employed throughout the

model's training phase. Therefore, the wallets dispersed over the internet may not be helpful for identifying the bit mixing services' new strategies since their underlying algorithms have changed. Thus, it is valuable to develop ML models trained on the most recent data.

In our thesis, we examined the Ethereum and Bitcoin transaction networks to detect illegal cryptocurrency transactions utilizing machine learning models such as logistic regression, random forest, and XGBoost. We generated features depending on the egonet networks and being independent of the networks. Additionally, we separated bitcoin wallets into two groups: ones that belong to Bitcoin mixing services and ones that do not belong to a specific illegal group such as Ponzi, gambling, or Bitcoin mixing services.

3. DATA COLLECTION AND TAGGING

We collected Bitcoin wallets from different resources such as bitcoinabuse.com, and walletexplorer.com. bitcoinabuse.com has revealed more than 200.000 illicit wallets, most of which are used for blackmail and fraud. Approximately 4000 new wallets are added to the list each month. In addition to wallets using Bitcoinfog and Helix services from walletexplorer.com, we also extracted wallets using legitimate KYC exchange services from this website. We utilized the Ethereum licit−illicit wallets dataset which was studied by Farrugia et al. [17]. After collecting all data, we used the following labeling rule while classifying Bitcoin data (Ethereum is already labeled.).

- ➢ If a wallet uses the known illegal Bitcoin mixing services such as BitcoinFog, and Helix, or it belongs to the list pulled by bitcoinabuse.com, then we tag it as an illicit wallet.
- ➢ If a wallet uses trustful cryptocurrency services using the Know-Your-Customer policy such as Binance, then we tag it as a licit wallet.

We defined five wallet datasets based on the previously mentioned rule: Bitcoin Illicit Wallets Dataset (B-Illicit), Bitcoin Licit Wallets Dataset (B-Licit), Helix Illicit Wallets Dataset (H-Illicit), BitcoinFog Illicit Wallets Dataset (BFog-Illicit), Ethereum Licit−illicit Wallets Dataset (Eth-Licit−illicit). As previously indicated, we retrieved the Ethereum data from the Farrugia et al. study [17]. The Ethereum community has identified this dataset as both licit and illicit. This dataset has 4681 wallets in total. 2502 of these wallets are legitimate, while 2179 are illegal. 3329 out of 4681 wallets' transactions were pulled in this study. Of the 3329 wallets, 1414 are illegal, while the remaining 1915 are legal. More than 70% of the wallets in these five wallet datasets started making their first transactions between 2017 and 2021 when initial transaction dates for the wallets are taken into account. These datasets are explained below in Table 3.1.

Table 3.1. Dataset Description.

Dataset Name	Source	Wallet Count
BIllicit	bitcoinabuse.com	5499
BLicit	KYC exchange services-walletexplorer.com	10000
HIllicit	Helix mixing service-walletexplorer.com	38704
BFogIllicit	BitFog mixing service-walletexplorer.com	10000
Eth-Licit−illicit	found on Kaggle	3329

Each historical transaction data was pulled using the Python blockchain API. In the data extraction process, we pulled the available historical transactions related to suspicious wallets. We created transaction networks by extracting the information of all the transactions of the suspicious wallets and the information of the other wallets that the suspicious have transacted. A hash is a mathematical function that converts an input of arbitrary length into an encrypted output of a fixed length [53]. As the input data changes, the outcome of the hash function changes too. In cryptocurrency terminology, a transaction is a set of transfers that have more than one input/output wallets. With each transaction, a unique hash is generated. In each Bitcoin transaction, 1-1, 1-N, N-1, and N-N relationships can be seen between input and output wallets. Also, each Bitcoin transaction can have the same input wallet or the same output wallet more than once. For Ethereum transactions, there is only one kind of relationship which is 1-1 (see in Figure 1). Therefore, the structure of the Ethereum network is simpler than the Bitcoin network.

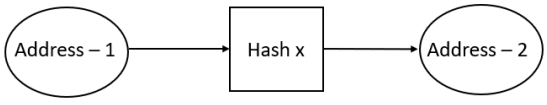

Figure 3.1. Ethereum transaction x with 1 input wallet and 1 output wallet.

After pulling historical transactions of wallets, we defined the Bitcoin and Ethereum transaction network graph as follows. In Bitcoin transactions where we have N-N relationships ($N > 1$), it is not always possible to determine how much cryptocurrency is sent from which input wallet to which output wallet easily. Therefore, we define two types of nodes in our network: hashes and wallets. If wallet - 1 is an input wallet for Hash x and wallet - 2 is an output wallet for Hash x, then we add an arc from wallet - 1 to Hash x and another arc from Hash x to wallet - 2. Each arc is weighted by the values of the corresponding input or output amount. An example of a network corresponding to a network that includes 3 different transactions can be seen in Figure-2. In this figure, we can see that a wallet can be found as input more than one. Also, an output wallet can be an input wallet for another transaction.

Table 3.2. Transaction Network Information.

Dataset Name	**Number Of Nodes**	**Number Of Edges**
BLicit	1065638	7310260
HIllicit	554266	2080662
BFogIllicit	432805	4001332

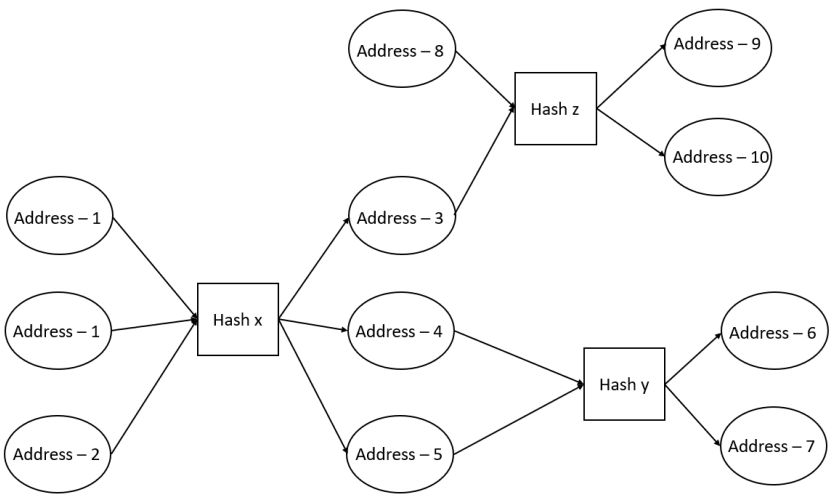

Figure 3.2. A Bitcoin network with 3 different transactions.

4. FEATURE EXTRACTION

In this part, we defined two different feature sets: egonet dependent features (EDF) and egonet idependent features (EDF).

4.1. Egonet Dependent Features Selection

We first tried to create graph-related features such as density, number of nodes and edges, average shortest path length, diameter, transitivity, average clustering coefficient, and number of triangles. We calculated these parameters for each components. Ego networks implies that each subgraph generated comprises of the $n-step-neighborhood$ as well as the connections between the central node's neighbors, which is referred as the central node's ego network. In general, the ego network of the i^{th} node is described as the subgraph covered by all of the i^{th} node's neighbors [54]. The set of $1-step-neighbors$ of a node i is made up of all nodes j that have a directed edge from i to j or a directed edge from j to i. In this study, we assumed $n = 2$ while generating the egonets of the wallets. As a result, the radius of the ego network is increased by one such that the neighbors of $1-step-neighbors$ who are not in direct interaction with the center node are included in the node set. This is because while the first neighbor of the wallet gives us the hashes its transactions contain, the second neighbor of a wallet also brings the information of other wallets in that transaction. These are commonly used egonet dependent network features (EDF):

- Density: The density calculates the degree of connectivity in a graph by dividing the number of edges by the number of all possible edges [55]. Assuming V is the total number of nodes and E is the total number of edges in a graph, the formula for calculating the density D of a directed graph is

$$\frac{2E}{V(V-1)}.$$

- Number of Nodes: The basic measure for a subgraph's volume is the sum of its member nodes. Each subgraph definitely has its own peculiarity.
- Number of Edges: The total number of edges in a subgraph represents the total number of transactions that occurred inside the subgraph. It is known as network size in short.
- Average Shortest Path Length: The geodesic distance, also known as the shortest path length between two nodes u and v, is the smallest distance value among all walks connecting u and v [56].
- Diameter: In networks, diameter is categorized among the distance and path measurements [55]. It is also known as the longest geodesic or the maximum eccentricity of a network, where eccentricity of a node u is defined as the greatest distance between u and any other node v in the network such that $v \in V$. When all feasible pairs are investigated, diameter represents the largest distance between two nodes.
- Average Clustering coefficient: Clustering coefficient is a structural metric that assesses the likelihood of nodes in a network to cluster together. In other words, it illustrates how near a node's neighbors are to constructing a complete network with unit density [57]. Let $|\mathrm{e}_n|$ be the number of edges having both ends in node n's $1 - step - neighborhood$, and k_n its degree [58]. The clustering coefficient CC_n of a node n is

$$\frac{2|e_n|}{k_n(k_n-1)}.$$

As a result, the average clustering coefficient CC of a network is determined by averaging the clustering coefficients of all nodes in the network, where V is the set of nodes

$$\frac{1}{|V|}\sum_{n\in V} CC_n.$$

- Number of Triangles: This feature, as the name implies, is the number of triangles present in the subgraph where all three nodes and edges are members of the same subgraph.
- Transitivity: The ratio of all potential triangles in a network equals transitivity. This characteristic has a similar logic to density measure; we describe it as the triangle realization ratio. Essentially, the number of triangles in the candidate subgraphs is divided by the maximum number of triangles, i. e. all feasible configurations comprising three nodes. This function, as the name suggests, displays network transmittance capacity.

Egonet structures, are made up of two different kinds of nodes: hash and wallets. There is no direct arc connecting two wallets since each transaction includes a hash in the middle. Due to the structure of the transaction graph, transitivity, average clustering coefficient, and the value of the number of triangles do not add any value to our model. On the other hand, density, number of nodes and edges in egonet graph, average shortest path length, and diameter are calculated.

4.2. Egonet Independent Features Selection

We generated other features in addition to egonet-dependent features. We are inspired by Li et al.'s study to generate some of these features [37]. The following list is showing an explanation of features that are independent from egonet networks (EIF). Each feature is calculated separately for the cases where an address is output and input. For instance, activity days, which is in the first place in the list below, is calculated separately for transactions where a wallet is an sender and for cases where it is an receiver.

- Activity Days: Number of transaction days for a wallet.
- Average Frequency Of Transaction Days: Average days between two consecutive transactions of a wallet on different days.
- Daily Maximum Transaction Count: The maximum number of transactions made

by a wallet as an input wallet in a day.

- Daily Minimum Transaction Count: The minimum number of transactions made by a wallet as an input wallet in a day.
- Total Transaction Count: Total number of transactions where a wallet is an input wallet.
- Minimum Frequency Of Transaction Days: Minimum day difference between two consecutive transactions of a wallet on different days.
- Maximum Input Wallet Count: Maximum number of other input wallets in transactions where one wallet is the input wallet.
- Maximum Input Value: The maximum cryptocurrency value a wallet sends in all transactions it is in.
- Maximum Output Wallet Count: Maximum number of other output wallets in transactions where one wallet is the output wallet.
- Maximum Output Value: The maximum cryptocurrency value a wallet receives in all transactions it is in.
- Minimum Input Wallet Count : Minimum number of other input wallets in transactions where one wallet is the input wallet.
- Minimum Input Value: The minimum cryptocurrency value a wallet sends in all transactions it is in.
- Minimum Output Wallet Count: Minimum number of other output wallets in transactions where one wallet is the output wallet.
- Minimum Output Value: The minimum cryptocurrency value a wallet receives in all transactions it is in.
- Sum Input Value: The total cryptocurrency value sent by a wallet in all transactions it is in.
- Sum Output Value: The total cryptocurrency value receive by a wallet in all transactions it is in.
- Average Input Value: The average cryptocurrency value sent by a wallet in all transactions it is in.
- Average Input Wallet Count : Average number of other input wallets in transactions where one wallet is the input wallet.

- Average Output Value: The average cryptocurrency value receive by a wallet in all transactions it is in.
- Average Output Wallet Count : Average number of other output wallets in transactions where one wallet is the output wallet.

5. EXPERIMENTAL RESULTS

5.1. Machine Learning

The topic of identifying illicit wallets in a cryptocurrency network is defined as a supervised binary classification problem, with positive observations representing illicit wallets and negative observations representing licit wallets. Classification models are primarily meant to generate probabilities for each observation about their membership to each class, and then allocate those probabilities to discrete classes in relation to a given threshold [59]. For our binary classification problems, we recommend using three methods: Random Forest, XGBoost, and Logistic Regression.

Random forest is a classification algorithm that consists of many independent decision trees with branches based on feature and function values in the leaf nodes. These many distinct decision trees perform together as an ensemble. The random forest's various trees each spit out a class prediction, and the class that receives the most votes becomes the prediction made by the model. The main principle of random forest is the wisdom of crowds since there are numerous uncorrelated models (trees) working as a committee. The most important factor is the low correlation between models. Using two techniques, random forest makes sure that the behavior of each individual tree is not overly connected. These methods are bagging (bootstrap aggregation) and feature randomness. The use of bagging in random forest is made possible by allowing each tree to randomly sample from the dataset with replacement, resulting in diverse trees. A random forest can only select one random subset of features per tree due to feature randomization. When splitting a node in a typical decision tree, we analyze all potential features and choose the one that results in the greatest gap between the observations in the left node and those in the right node. We carried out hyperparameter tinkering to optimize the random forest classifier throughout a parameter space using a 5-fold cross-validation method. Three parameters have been chosen for tuning: maximum depth of three, minimum sample split, and maximum leaf nodes. The maximum

depth of a tree in Random Forest is the longest path from the root node to the leaf node. In a random forest, the decision tree is instructed by the function min sample split to split any node that has less observations than necessary. Max leaf nodes limits the growth of the tree by imposing a restriction on the splitting of the nodes in the tree.

Gradient boosted decision trees are implemented in a high-performance manner by the open source package known as XGBoost. A single model is trained on the dataset and used for prediction with a typical machine learning model, like a decision tree. The data can be expanded or the parameters modified, but only one model is used in the end. Even if an ensemble is produced, each model is trained separately and used on the data. Contrarily, boosting employs a more iterative methodology. Models are gradually introduced till no more advancements are possible. This iterative technique has the advantage that new models added focus on resolving errors caused by older models. Models trained individually in a conventional ensemble technique may all commit the same errors. To fine-tune the parameters in XGboost, we used 5-fold stratified cross-validation, just like in Random forest. Learning rate and max depth are tuned. An optimization algorithm's learning rate controls the step size at each iteration as it advances toward the least loss function.

A predictive analytic algorithm based on the idea of probability, logistic regression is a machine learning approach that is used for categorization problems. The cost function employed in logistic regression is more sophisticated and is referred to as the sigmoid function. The sigmoid function is used to match the predicted values to the probability. The sigmoid function is usually restricted to the range of 0 and 1 according to the logistic regression hypothesis. newton-cg Solver is used. There are no necessary hyperparameters to adjust in logistic regression.

5.2. Model Setup

We have 5 fundamental datasets for each wallet datesets as indicated in Table 5.1. We have one licit wallet dataset and three illicit datasets for Bitcoin. At the end of the feature selection, we obtained four feature datasets for each dataset indicated in Table 5.1 by utilizing their transaction network pulled by Python API. While generating feature datasets, we merged the Blicit wallets' feature dataset and other Bitcoin illicit wallet feature datasets. After doing that, we defined 4 datasets for training models.

Table 5.1. Feature Datasets Combination.

Dataset Name	**Alias**
BIllicit−BLicit	Bitcoin
HIllicit−BLicit	Helix
BFogIllicit−BLicit	Bitcoinfog
Eth-Licit−illicit	Ethereum

We accumulated 31 features for Bitcoin, Helix and Bitcoinfog datasets and 26 features for the Ethereum dataset to detect suspicious wallets using machinel learning models such as logistic regression(LR), random forest(RF), and XGBoost(XGB). Egonet-dependent features are only calculated for the Bitcoin network. Due to Ethereum transaction nature, egonet dependent features are proportional to some egonet independent features. For instance, density, the number of nodes, and the number of edges are proportional to the total hash count. However, other features that are not included in the egonet features are common for both Ethereum and Bitcoin networks.

5.2.1. Hyperparameter Tuning

Supervised binary classification models are used to detect suspicious financial behaviors in cryptocurrency transaction networks. First of all, we divided the datasets using an 80:20 split ratio. In other words, 80% of the wallets in the datasets are used as the training set, while 20% of them are used as the testing set. For each

dataset, we trained three classifiers with three different combinations of feature sets. Furthermore, we performed 5 stratified folds cross-validation on the training set to tune the parameters of classifiers. We tested our proposed system with four different datasets and two different binary classifiers: random forest(RF), and XGBoost(XGB). We used a laptop with Intel Core i7 - 7500U CPU and 8GB of memory. Tested parameters for random forest max depth: (4,8), min samples leaf :(2,4), min samples split: (2,10), n estimators: (1000) and for XGBoost max depth:(4,8), learning rate: (0.1,0.2), n estimators: (1000). While comparing the performance of different models, we used AUC values. The results of our cross-validations are reported in Appendix A. The best performing classifiers with hyperparameters and feature sets are presented in Table 5.2. The dataset column in this table shows from which datasets illicit and licit wallets were drawn. EIF indicates that the classifier above the name of the classifier works only with egonet-independent features.

Table 5.2. Best-performing model settings.

Dataset	**ML Model**	*learning_rate*	*max_depth*	**AUC**
Bitcoinfog	XGB^{EIF}	0.2	4	0,999993
Helix	XGB^{EIF}	0.1	4	0.999999
Bitcoin	XGB^{EIF}	0.1	4	0.999991
Ethereum	XGB^{EIF}	0.1	4	0.956237

5.3. Results

5.3.1. Evaluation Metrics

We assess classifiers' performance using standard machine learning metrics like recall, precision, the area under the receiver operating characteristic curve (AUC), and f2 score. Recall, also named for sensitivity is the percentage of true positives that the machine learning model properly detects. Only the classification of the positive

instances is significant to the recall [60]. The documentation of the confusion matrix contains the definitions of true positive and false negative used in the formulation below. The documentation states that when both the actual classification and the predicted classification are positive, it is known as a true positive (TP) result since the classifier successfully detected the positive sample. On the other hand, a false negative (FN) result is produced when the actual classification is positive but the predicted classification is negative. This occurs when the classifier misclassifies the positive sample as being negative [61]. Recall is

$$\frac{TP}{TP + FN}.$$

Precision means that the ratio of properly diagnosed positive samples to all samples that were classified as positive is used to calculate the precision (either correctly or incorrectly). In the Precision formula, a false positive (FP) result is one in which the predicted classification is positive but the actual classification is negative. This occurs when the classifier wrongly classifies the negative sample as positive [61]. Precision is

$$\frac{TP}{TP + FP}.$$

Huang et al. suggest that since AUC is a more accurate metric, it should take the role of accuracy when evaluating and comparing classifiers [62]. Therefore, the area under the receiver operating characteristic curve (ROC) is often used as an important evaluation statistic for classifier problems. On the Y and X axes of a ROC graph, respectively, true positive rates (TPR) and false positive rates (FPR) are shown. The true positive rate is equal to the recall. On the other hand, the ratio between outcomes where the model mistakenly predicts the negative class and all actual negatives is known as the false positive rate. Each classifier in the ROC space is represented by a point (FPR, TPR) on the ROC curve.TPR and FPR may vary (0 and 1) as the output threshold for a model that produces output continuously varies between its endpoints. ROC curve is the name of the resulting curve. The abbreviation “Area under the ROC Curve” is AUC. In other words, AUC measures the complete two-dimensional region beneath the

entire ROC curve from (0,0) to (1,1) [62]. The harmonic mean of recall and precision is used to get the F1-score [63]. F1-score is

$$\frac{2 * Precision * Recall}{Precision + Recall}.$$

5.3.2. Experimental Results

We used the hyperparameters indicated in Table 5.2 to evaluate performance of our models on the test data. The outcomes for the best performing models employing all datasets are shown in Table 5.3. Since logistic regression does not have any hyper-parameter, we could not evaluate its performance in cross-validation process. Instead, we trained a logistic regression model on the train dataset and tested its performance on the test dataset. The results of the logistic regression models are also displayed in this table together with the best-performing models in Table 5.3.

Table 5.3. Performances on test set.

Exp.	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	XGB^{EIF}	0.999248	0.999252	0.999249	0.999987
2	**Helix**	XGB^{EIF}	0.999935	0.999754	0.999844	0.999999
3	**Bitcoin**	XGB^{EIF}	0.997894	0.998547	0.998219	0.999986
4	**Ethereum**	XGB^{EIF}	0.892071	0.895582	0.893651	0.962910
5	**Bitcoinfog**	LR^{EIF}	0.968038	0.967714	0.967742	0.990970
6	**Helix**	LR^{EIF}	0.993998	0.990468	0.992221	0.999183
7	**Bitcoin**	LR^{EIF}	0.954656	0.941948	0.947712	0.972486
8	**Ethereum**	LR^{EIF}	0.708722	0.691764	0.695328	0.813703

5.3.3. Class Imbalance Challenges in Real-life Application

There are several viewpoints regarding the ratio of licit to illicit wallets in real life. Although there is a consensus on the imbalance of illicit/licit wallets in real life, there is no agreement on the imbalance ratio. In 2021, 0.15% of transactions were deemed suspicious by Chainanalysis, an American blockchain analysis company founded in 2014 that sells tools to assist clients with blockchain analysis. Sean Foley questioned Chainanalysis in an essay titled Sex, Drugs, and Bitcoin that was published in Review of Financial Studies in 2019 and said that while 0.46% of transactions are suspicious, only 25% of bitcoin users are. The Financial Action Task Force, the international organization responsible for advancing money laundering norms for traditional finance and the crypto industry, reports that suspicious transaction rates, given between 0.1% and 15.4%, are extremely low. According to Coindesk's research, estimates of the size of online crime range from less than 1% to around 50% of all crypto activity [64]. As a result, we repeated our studies with different ratios.

Table 5.4. Performances on Bitcoinfog test set with different illicit wallet ratio.

Ratio	**Dataset**	**ML Model**	**Precision**	**Recall**	**F1 score**	**AUC**
0,01	**Bitcoinfog**	XGB^{EIF}	0,966167	0,937250	0,951238	0,997738
0,02	**Bitcoinfog**	XGB^{EIF}	0,987844	0,987844	0,987844	0,999976
0,10	**Bitcoinfog**	XGB^{EIF}	0,999504	0,994505	0,996989	0,999991
0,50	**Bitcoinfog**	XGB^{EIF}	0.999248	0.999252	0.999249	0.999987

Table 5.5. Performances on Helix test set with different illicit/licit wallet ratio.

Ratio	**Dataset**	**ML Model**	**Precision**	**Recall**	**F1 score**	**AUC**
0,01	**Helix**	XGB^{EIF}	0.970588	0.999750	0.984723	0.999578
0,02	**Helix**	XGB^{EIF}	0.987844	0.987844	0.987844	0.999976
0,10	**Helix**	XGB^{EIF}	0.999752	0.997252	0.998498	0.999624
0,60	**Helix**	XGB^{EIF}	0.999935	0.999754	0.999844	0.999999

Table 5.6. Performances on Bitcoin test set with different illicit/licit wallet ratio.

Ratio	**Dataset**	**ML Model**	**Precision**	**Recall**	**F1 score**	**AUC**
0,01	**Bitcoin**	XGB^{EIF}	0.966167	0.937250	0.951238	0.999875
0,02	**Bitcoin**	XGB^{EIF}	0.974	0.951880	0.962664	0.999761
0,10	**Bitcoin**	XGB^{EIF}	0.999752	0.997252	0.998498	0.999997
0,35	**Bitcoin**	XGB^{EIF}	0.997894	0.998547	0.998219	0.999986

Table 5.7. Performances on Ethereum test set with different illicit/licit wallet ratio.

Ratio	**Dataset**	**ML Model**	**Precision**	**Recall**	**F1 score**	**AUC**
0,02	**Ethereum**	XGB^{EIF}	0.881035	0.766585	0.812918	0.860398
0,05	**Ethereum**	XGB^{EIF}	0.910751	0.799603	0.845004	0.928476
0,15	**Ethereum**	XGB^{EIF}	0.849510	0.838343	0.843769	0.951044
0,40	**Ethereum**	XGB^{EIF}	0.892071	0.895582	0.893651	0.962910

5.3.4. Result Discussion

XGBoost outperformed random forest and logistic regression on all datasets. Judging from the cross-validation results, we conclude that the performance of models on Bitcoin datasets is less sensitive to hyperparameter selections. However, this is not true for the Ethereum dataset. Also, when we adjust the test datasets so that there are different degrees of class imbalances; the performance of XGB on Bitcoin datasets did not degrade but on the Ethereum dataset we experienced a significant decrease. One potential root cause of the sensitivity of models trained on the Ethereum dataset is that class labels are more accurate than that of Bitcoin datasets due to the fact that there are no specific rules for gathering licit wallets. We assumed that addresses using know-your-customer exchange services are licit which may lead to some degree of mislabelling. Egonet dependent features do not significantly improve models. Hence, we have used egonet independent features considering computational complexity.

5.3.4.1. Detecting Suspicious Bitcoin Mixer Wallets. In this section, we examined the results of the models which use BFogIllicit−Licit dataset based on Table 5.3. The AUC of experiment-3 is 99.99%. The illicit ratio of this experiment is 50%.

On the other hand, we see that models using HIllicit−Licit dataset outperform models using BFogIllicit−Licit dataset in terms of precision, recall, f1 score and AUC in Table 5.2. Therefore, HIllicit−Licit dataset makes it simpler to distinguish between licit and illicit activity than BFogIllicit−Licit. The illicit ratio of this experiment is 60%.

5.3.4.2. Detecting Suspicious Bitcoin Wallets. Based on what we previously stated, we continue using the models with the EIF feature set in this part due to both the inadequate improvement and the computational complexity. Table 5.3 shows that the model utilizing LR yields the worst outcomes in accordance with the evaluation criterias. XGB performs better in terms of precision, recall, and f1 score. Although we

have obtained results close to HIllicit−Licit and BFogIllicit−Licit, there is a difference in features that affect the performance of the models. The illicit ratio of this experiment is 60%.

5.3.4.3. Detecting Suspicious Ethereum Wallets. As we mentioned before, there is just one type of connection for Ethereum transactions, and that is a 1-1 relationship (see in Figure 1). As a result, the Ethereum network's structure is less complex than the Bitcoin network's. Besides, despite the fact that there are instances in which a wallet is exclusively utilized as a receiver in Bitcoin transactions, we did not come across any examples of this case in Ethereum transactions.

We also observed that the estimated features of illegal and legal wallets are similar to each other in the Ethereum dataset. Thus, the EIF feature set is not as sufficient for Ethereum as it is for the Bitcoin dataset to distinguish between licit and illicit wallets. As a result, the results in Table 5.3 are less promising than those obtained with models that used datasets containing Bitcoin wallets. AUC is over 99% for models using Bitcoin datasets whereas AUC is 96% for Ethereum dataset. The illicit ratio of this experiment is 38%. In conclusion, we cannot use the same feature set for models using Ethereum to obtain the same performance in models using Bitcoin since the data structures for Ethereum and Bitcoin are different.

5.3.4.4. Feature Importance Analysis. Feature importance is extracted for the best models presented in Table 5.3. As described in Section 4.2, in each dataset during feature engineering, two types of features regarding transactions a specific wallet involved are defined. One is when a wallet is the receiver and the other is when a wallet is the sender. In Figure 5.1, 5.2 and 5.3, _x represents the features generated regarding transactions in which a specific wallet is the sender, and _y represents otherwise. In Appendix B, histograms of important features through illicit and licit wallets for all datasets are provided. Note that in the histograms bins are stacked i.e. bins of illicit wallets are stacked on top of the bins of licit wallets.

We see that the min_input_value feature is the most important feature for Bitcoinfog and Helix dataset (see Figure 5.1 and 5.2). THe histogram of this feature on aforementioned datasets visually indicates how the distribution varies between instances of different classes i.e. illicit and licit wallets (see in Figure B.1 and B.2). The fact that only one feature has meaningful importance address that the class mislabelling issue in Section 5.3.4.

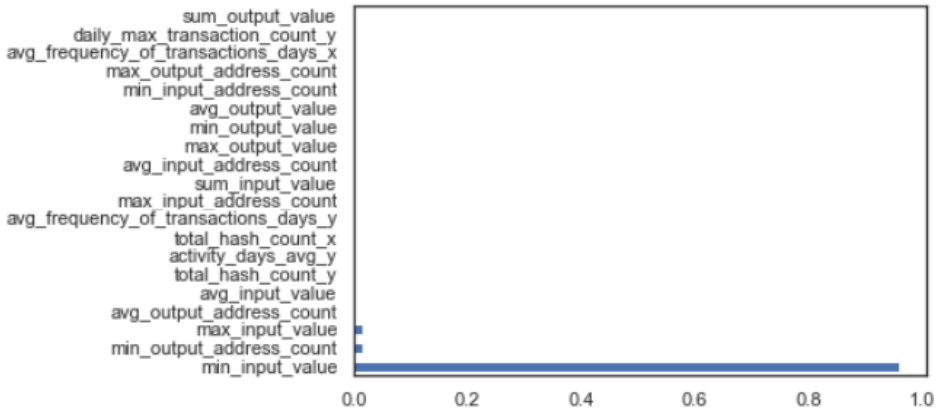

Figure 5.1. Feature importance - BFogIllicit−Licit.

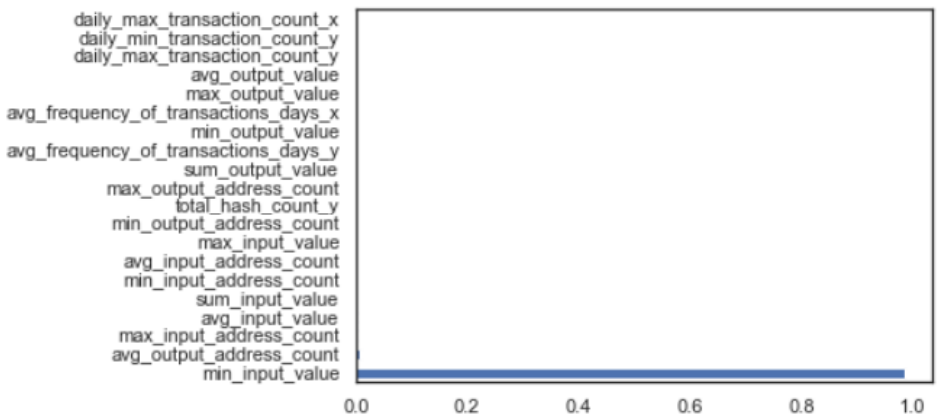

Figure 5.2. Feature importance - HIllicit−Licit.

Important features for Bitcoin dataset is different from Bitcoinfog and Helix datasets (see Figure 5.3). There are more output wallets in a typical licit Bitcoin transactions than that of a typical illicit Bitcoin transactions. Furthermore, illicit wallets often send smaller amounts of bitcoin than licit wallets do through lifetime.

Moreover, min_output_address_count denotes the minimum number of receiver wallet count among all transactions a particular wallet has involved, and approximately 85% of illicit Bitcoin wallets have less than 10 min_output_address_count, which is extremely distinctive (see in Figure B.3 and B.4).

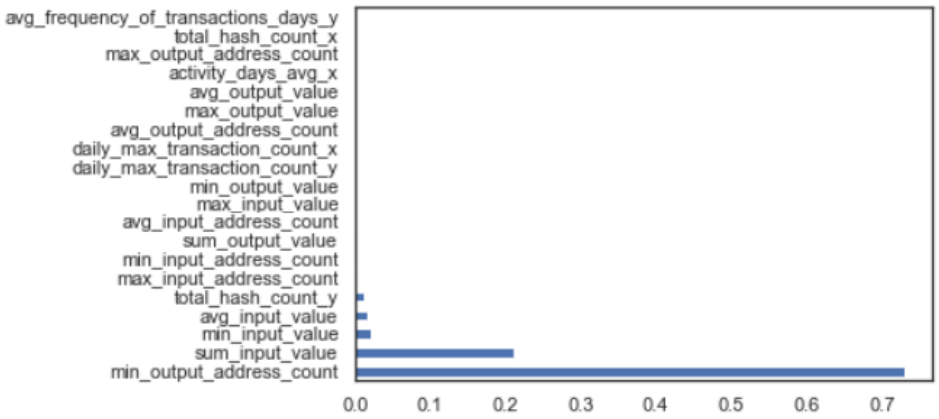

Figure 5.3. Feature importance - BIllicit−Licit.

When comparing important features of Ethereum dataset with those of other three datasets, we notice several differences. These also validate that Ethereum data structure differs from the rest. Illicit wallets make more transactions in a day and receive less ethereum than licit wallets on average and at minimum. We observe that the wallets having relatively long lifespan are licit.

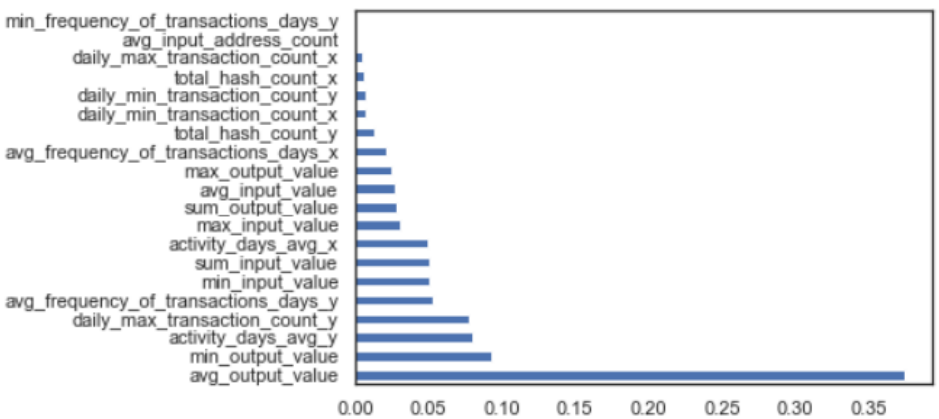

Figure 5.4. Feature importance - Eth−Licit−Illicit.

6. CONCLUSION

At the end of the study, we tested several models to detect suspicious wallets in different Ethereum and Bitcoin datasets described in Table 3.1. XGB outperfomed other classifiers on all of the datasets.

To begin with, we observed that suspicious wallets of datasets in which mixer services are utilized, are easier to detect. However, this situation might have arised due to the fact that these datasets might be suffering from mislabelling issue.

Furthermore, models' performance did not benefit sufficiently from egonet dependent features, and these features have a high computational complexity. Despite the fact that the computational power required to extract these features are not negligible.

Moreover, the performance of the models were not as promising as the models trained on Bitcoin datasets when we attempted to detect suspicious Ethereum wallets using the same egonet independent features. This is because Ethereum has a different transaction structure than Bitcoin.

Despite the fact that high AUC values appear too good to be true, it can be observed that these high values are also obtained in related researches in the literature [17, 18, 35, 42]. It should be kept in mind that when different levels of class imbalances are introduced to the datasets, the performances deteriorated.

The predicted suspicion probability of these models can be utilized to assign suspicion score to wallets. As a future scenario, when financial institutes such as banks ask their customers to share their cryptocurrency wallet, they will have another indicator to fight against money laundering.

REFERENCES

1. "Home : Oxford English Dictionary", `https://www.oed.com/`, accessed on May 17, 2022.

2. Chaum, D., "Blind Signatures For Untraceable Payments", *Advances In Cryptology*, pp. 199–203, 1983.

3. "Satoshi Nakamoto", `https://en.wikipedia.org/wiki/Satoshi_Nakamoto/`, accessed on December 3, 2022.

4. Muller, W. H., C. Kalin and J. G. Goldsmith, *Anti-money Laundering: International Law and Practice*, John Wiley & Sons / Henley & Partners, Chichester, West Sussex, England, 2007.

5. Trautman, L., "Virtual currencies: Bitcoin & What Now After Liberty Reserve, Silk Road, And Mt. Gox", *Rich. JL & Tech.*, Vol. 20, p. 1, 2013.

6. Meiklejohn, S., M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker and S. Savage, "A Fistful Of Bitcoins: Characterizing Payments Among Men With No Names", *Internet Measurement Conference*, pp. 127–140, 2013.

7. Nick, J. D., *Data-driven De-anonymization In Bitcoin*, Master's Thesis, ETH-Zürich, 2015.

8. Reid, F. and M. Harrigan, "An Analysis Of Anonymity In The Bitcoin System", *Security and Privacy In Social Networks*, pp. 197–223, Springer, 2013.

9. Hong, Y., H. Kwon, J. Lee and J. Hur, "A Practical De-Mixing Algorithm For Bitcoin Mixing Services", *Proceedings of the 2nd ACM Workshop on Blockchains, Cryptocurrencies, and Contracts*, pp. 15–20, 2018.

10. Harrigan, M. and C. Fretter, "The Unreasonable Effectiveness Of Address Clustering", *Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld)*, pp. 368–373, 2016.

11. Koshy, P., D. Koshy and P. McDaniel, "An Analysis Of Anonymity In Bitcoin Using P2P Network Traffic", *International Conference on Financial Cryptography and Data Security*, pp. 469–485, 2014.

12. Biryukov, A. and I. Pustogarov, "Bitcoin Over Tor Isn't A Good Idea", *Symposium on Security and Privacy*, pp. 122–134, 2015.

13. "Crystal Blockchain Report 2020", `https://crystalblockchain.com/articles-/crystalblockchain-end-of-year-report-2020/`, accessed on November 3, 2022.

14. Wu, L., Y. Hu, Y. Zhou, H. Wang, X. Luo, Z. Wang, F. Zhang and K. Ren, "Towards Understanding And Demystifying Bitcoin Mixing Services", *Proceedings of the Web Conference*, pp. 33–44, 2021.

15. Liu, C., J. Wang and M. Wang, "Toward Understanding the Relationship Between Bitcoin Mixing Services And Circular Transactions", *PACIS Proceedings*, 2022.

16. Kus Khalilov, M. C. and A. Levi, "A Survey on Anonymity and Privacy in Bitcoin-Like Digital Cash Systems", *IEEE Communications Surveys Tutorials*, Vol. 20, No. 3, pp. 2543–2585, 2018.

17. Farrugia, S., J. Ellul and G. Azzopardi, "Detection Of Illicit Accounts Over The Ethereum Blockchain", *Expert Systems With Applications*, Vol. 150, p. 11318, 2020.

18. Vassallo, D., V. Vella and J. Ellul, "Application Of Gradient Boosting Algorithms

for Anti-money Laundering In Cryptocurrencies", *SN Computer Science*, Vol. 2, p. 143, 2021.

19. Mukhopadhyay, U., A. Skjellum, O. Hambolu, J. Oakley, L. Yu and R. Brooks, "A Brief Survey Of Cryptocurrency Systems", *14th Annual Conference on Privacy, Security and Trust (PST)*, pp. 745–752, 2016.

20. "Blockchain", `https://en.wikipedia.org/wiki/Blockchain`, accessed on November 3, 2022.

21. "The Great Chain Of Being Sure About Things | The Economist", `https://web.archive.org/web/20160703000844/http://www.economist.com/-news/ briefing/21677228-technology-behind-bitcoin-lets-people-who-do--not-know-or-trust-each-other-build-dependable`, accessed on November 3, 2022.

22. Fauzi, M. A., N. Paiman and Z. Othman, "Bitcoin And Cryptocurrency: Challenges, Opportunities And Future Works", *The Journal of Asian Finance, Economics and Business*, Vol. 7, No. 8, pp. 695–704, 2020.

23. "Markets Price Charts", `https://markets.bitcoin.com/crypto/BTC/`, accessed on, September 4, 2022.

24. Böhme, R., N. Christin, B. Edelman and T. Moore, "Bitcoin: Economics, Technology, and Governance", *Journal of Economic Perspectives*, Vol. 29, No. 2, pp. 213–38, 2015.

25. "BitcoinFog All About Cryptocurrency", `https://en.bitcoinwiki.org/-wiki/Bitcoin_Fog`, accessed on September 4, 2022.

26. "Bitcoin Mixing - For Money Laundering", `https://www.businessinsider.in/-investment/news/bitcoin-mixing-service-helix-pleads-guilty-to-laundering-300-million-will- forfeit-over-4400-bitcoins/articleshow-`

/85482090.cms, accessed on, November 4, 2022.

27. Buterin, V. *et al.*, "A Next-Generation Smart Contract And Decentralized Application Platform", Vol. 3, No. 37, pp. 2–1, 2014.

28. Vujičić, D., D. Jagodić and S. Ranić, "Blockchain Technology, Bitcoin, And Ethereum: A Brief Overview", *17th International Symposium InfotehJahorina (Infoteh)*, pp. 1–6, 2018.

29. Chen, S., "Cryptocurrency Financial Risk Analysis Based on Deep Machine Learning", *Complexity*, 2022.

30. Alessandretti, L., A. ElBahrawy, L. M. Aiello and A. Baronchelli, "Anticipating Cryptocurrency Prices Using Machine Learning", *Complexity*, 2018.

31. Akyildirim, E., A. Goncu and A. Sensoy, "Prediction of Cryptocurrency Returns Using Machine Learning", *Annals of Operations Research*, Vol. 297, No. 1, pp. 3–36, 2021.

32. Koker, T. E. and D. Koutmos, "Cryptocurrency Trading Using Machine Learning", *Journal of Risk and Financial Management*, Vol. 13, No. 8, 2020.

33. Attanasio, G., L. Cagliero, P. Garza and E. Baralis, "Quantitative Cryptocurrency Trading: Exploring The Use Of Machine Learning Techniques", *Proceedings of the 5th Workshop on Data Science for Macro-modeling with Financial and Economic Datasets*, pp. 1–6, 2019.

34. Sattarov, O., A. Muminov, C. W. Lee, H. K. Kang, R. Oh, J. Ahn, H. J. Oh and H. S. Jeon, "Recommending Cryptocurrency Trading Points With Deep Reinforcement Learning Approach", *Applied Sciences*, Vol. 10, No. 4, p. 1506, 2020.

35. Aziz, R. M., M. F. Baluch, S. Patel and A. H. Ganie, "LGBM: A Machine Learning Approach For Ethereum Fraud Detection", *International Journal of Information*

Technology, 2022.

36. Hu, Y., S. Seneviratne, K. Thilakarathna, K. Fukuda and A. Seneviratne, "Characterizing and Detecting Money Laundering Activities on the Bitcoin Network", *arXiv:1912.12060 [cs]*, 2019.

37. Li, Y., Y. Cai, H. Tian, G. Xue and Z. Zheng, "Identifying Illicit Addresses In Bitcoin Network", *International Conference on Blockchain and Trustworthy Systems*, pp. 99–111, 2020.

38. Lee, C., S. Maharjan, K. Ko, J. Woo and J. W.-K. Hong, "Machine Learning Based Bitcoin Address Classification", *International Conference on Blockchain and Trustworthy Systems*, pp. 517–531, Springer, 2020.

39. Chen, B., F. Wei and C. Gu, "Bitcoin Theft Detection Based on Supervised Machine Learning Algorithms", *Security and Communication Networks*, Vol. 2021, 2021.

40. Camino, R. D., R. State, L. Montero and P. Valtchev, "Finding Suspicious Activities in Financial Transactions and Distributed Ledgers", *International Conference on Data Mining Workshops (ICDMW)*, 2017.

41. Nerurkar, P., Y. Busnel, R. Ludinard, K. Shah, S. Bhirud and D. Patel, "Detecting Illicit Entities In Bitcoin Using Supervised Learning Of Ensemble Decision Trees", *10th international conference on information communication and management*, 2020.

42. Alarab, I., S. Prakoonwit and M. I. Nacer, "Comparative Analysis Using Supervised Learning Methods For Anti-Money Laundering In Bitcoin", *5th International Conference on Machine Learning Technologies*, pp. 11–17, 2020.

43. Nerurkar, P., S. Bhirud, D. Patel, R. Ludinard, Y. Busnel and S. Kumari, "Supervised Learning Model For Identifying Illegal Activities In Bitcoin", *Applied*

Intelligence, Vol. 51, No. 6, pp. 3824–3843, 2021.

44. Wu, J., J. Liu, W. Chen, H. Huang, Z. Zheng and Y. Zhang, "Detecting Mixing Services via Mining Bitcoin Transaction Network With Hybrid Motifs", *Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 52, No. 4, pp. 2237–2249, 2022.

45. Maksutov, A. A., M. S. Alexeev, N. O. Fedorova and D. A. Andreev, "Detection of Blockchain Transactions Used in Blockchain Mixer of Coin Join Type", *Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, pp. 274–277, 2019.

46. Möser, M., R. Böhme and D. Breuker, "An Inquiry Into Money Laundering Tools In The Bitcoin Ecosystem", *APWG eCrime Researchers Summit*, 2013.

47. Sun Yin, H. H., K. Langenheldt, M. Harlev, R. R. Mukkamala and R. Vatrapu, "Regulating Cryptocurrencies: A Supervised Machine Learning Approach To De-Anonymizing The Bitcoin Blockchain", *Journal of Management Information Systems*, Vol. 36, No. 1, pp. 37–73, 2019.

48. Jung, E., M. Le Tilly, A. Gehani and Y. Ge, "Data Mining-Based Ethereum Fraud Detection", *International Conference on Blockchain (Blockchain)*, pp. 266–273, 2019.

49. Li, M., Q. Wang and W. Zhang, "Blockchain-Based Distributed Machine Learning Towards Statistical Challenges", *International Conference on Blockchain and Trustworthy Systems*, pp. 549–564, 2020.

50. Awasthi, A., *Clustering Algorithms For Anti-Money Laundering Using Graph Theory And Social Network Analysis*, Centre de Recerca Matemàtica, 2012.

51. Soltani, R., U. T. Nguyen, Y. Yang, M. Faghani, A. Yagoub and A. An, "A New Algorithm For Money Laundering Detection Based On Structural Similarity",

7th Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON), pp. 1–7, 2016.

52. Noble, C. C. and D. J. Cook, "Graph-based Anomaly Detection", *9th ACM SIGKDD International Conference On Knowledge Discovery and Data Mining*, pp. 631–636, 2003.

53. "Understanding Hash", `https://www.investopedia.com/terms/h/hash.asp`, accessed on September 4, 2022.

54. Crossley, N., E. Bellotti, G. Edwards, M. G. Everett, J. Koskinen and M. Tranmer, *Social Network Analysis For Ego-Nets: Social Network Analysis For Actor-Centred Networks*, Sage, 2015.

55. Silva, T. C. and L. Zhao, "Complex Networks", *Machine Learning in Complex Networks*, pp. 15–70, 2016.

56. Dijkstra, E. W., "A Note On Two Problems In Connexion With Graphs", *Numerische Mathematik*, Vol. 1, No. 1, pp. 269–271, 1959.

57. Watts, D. J. and S. H. Strogatz, "Collective Dynamics Of Small-World Networks", Vol. 393, No. 6684, pp. 440–442, 1998.

58. Taş, G., *Network-Based Methods For Anti-Money Laundering*, Master's Thesis, Boğaziçi University, 2020.

59. Kuhn M., J. K., *Applied Predictive Modeling*, Vol. 26, Springer, 2013.

60. Altman, D. G. and J. M. Bland, "Diagnostic tests. 1: Sensitivity and Specificity.", *BMJ: British Medical Journal*, Vol. 308, No. 6943, p. 1552, 1994.

61. "sklearn.metrics.confusion_matrix", `https://scikit-learn/stable/modules-/generated/sklearn.metrics.confusion_matrix.html`, accessed on September

2, 2022.

62. Huang, J. and C. X. Ling, "Using AUC and Accuracy In Evaluating Learning Algorithms", *Transactions On Knowledge And Data Engineering*, Vol. 17, No. 3, pp. 299–310, 2005.

63. "F-score", `https://en.wikipedia.org/w/index.php?title=F-scoreoldid-=1106772144`, accessed on September 2, 2022.

64. "How Big Is Crypto Crime?", `https://www.coindesk.com/policy/2022/05/09-/how-big-is-crypto-crime-really/`, accessed on October 10, 2022.

65. Balthasar, T. d. and J. Hernandez-Castro, "An Analysis Of Bitcoin Laundry Services", *Nordic Conference on Secure IT Systems*, pp. 297–312, 2017.

66. Fanusie, Y. J. and T. Robinson, "Bitcoin Laundering: An Analysis Of Illicit Flows Into Digital Currency Services", *Center on Sanctions and Illicit Finance Elliptic*, 2018.

67. Bartoletti, M., B. Pes and S. Serusi, "Data Mining For Detecting Bitcoin Ponzi Schemes", *Crypto Valley Conference On Blockchain Technology (CVCBT)*, pp. 75–84, 2018.

68. Needham, M. and A. E. Hodler, *Graph Algorithms: Practical Examples In Apache Spark and Neo4j*, O'Reilly Media, 2019.

APPENDIX A: CROSS VALIDATION RESULTS

Table A.1. RF-Bitcoinfog (max_depth:4, min_samples_split:2, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998438	0,998438	0,998437	0,999969
2	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997503	0,997499	0,997499	0,999929
3	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998439	0,998436	0,998437	0,999941
4	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,995943	0,995936	0,995937	0,999836
5	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998438	0,998437	0,998437	0,999980
Avg	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997752	0,997749	0,997749	0,999931

Table A.2. RF-Bitcoinfog (max_depth:8, min_samples_split:2, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,999374	0,999374	0,999374	0,999994
2	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998128	0,998124	0,998124	0,999982
3	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998753	0,998749	0,998749	0,999995
4	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,996888	0,996873	0,996874	0,999961
5	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,999374	0,999374	0,999374	0,999995
Avg	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998504	0,998499	0,998499	0,999985

Table A.3. RF-Bitcoinfog (max_depth:4, min_samples_split:10, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998127	0,998125	0,998124	0,999969
2	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997503	0,997499	0,997499	0,999928
3	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998439	0,998436	0,998437	0,999941
4	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,995943	0,995936	0,995937	0,999837
5	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998438	0,998437	0,998437	0,999979
Avg	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997690	0,997687	0,997687	0,999931

Table A.4. RF-Bitcoinfog (max_depth:8, min_samples_split:10, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,999374	0,999374	0,999374	0,999993
2	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998128	0,998124	0,998124	0,999982
3	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998439	0,998436	0,998437	0,999994
4	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,996888	0,996873	0,996874	0,999956
5	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998749	0,998749	0,998749	0,999996
Avg	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998316	0,998311	0,998312	0,999984

Table A.5. RF-Bitcoinfog (max_depth:4, min_samples_split:2, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998438	0,998438	0,998437	0,999970
2	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997503	0,997499	0,997499	0,999930
3	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998128	0,998124	0,998124	0,999941
4	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,995943	0,995936	0,995937	0,999830
5	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998438	0,998437	0,998437	0,999979
Avg	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997690	0,997687	0,997687	0,999930

Table A.6. RF-Bitcoinfog (max_depth:8, min_samples_split:2, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998438	0,998438	0,998437	0,999970
2	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997503	0,997499	0,997499	0,999930
3	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998128	0,998124	0,998124	0,999941
4	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,995943	0,995936	0,995937	0,999830
5	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998438	0,998437	0,998437	0,999979
Avg	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997690	0,997687	0,997687	0,999930

Table A.7. RF-Bitcoinfog (max_depth:4, min_samples_split:10, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998127	0,998125	0,998124	0,999969
2	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997503	0,997499	0,997499	0,999930
3	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998128	0,998124	0,998124	0,999939
4	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,995943	0,995936	0,995937	0,999833
5	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998124	0,998124	0,998124	0,999979
Avg	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997565	0,997562	0,997562	0,999930

Table A.8. RF-Bitcoinfog (max_depth:8, min_samples_split:10, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,999374	0,999374	0,999374	0,999993
2	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,997814	0,997811	0,997812	0,999982
3	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998439	0,998436	0,998437	0,999992
4	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,996888	0,996873	0,996874	0,999951
5	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998749	0,998749	0,998749	0,999995
Avg	**Bitcoinfog**	$RF^{\text{EIF.}}$	0,998253	0,998249	0,998249	0,999983

Table A.9. RF model results Using Bitcoinfog (max_depth:4, min_samples_split:2, min_samples_leaf:2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996563	0,996563	0,996562	0,999939
2	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996258	0,996248	0,996249	0,999882
3	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,997198	0,997186	0,997187	0,999918
4	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,995314	0,995311	0,995312	0,999756
5	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996565	0,996561	0,996562	0,999947
Avg	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996380	0,996374	0,996374	0,999888

Table A.10. RF-Bitcoinfog (max_depth:8, min_samples_split:2, min_samples_leaf:2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,999062	0,999062	0,999062	0,999991
2	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998128	0,998124	0,998124	0,999976
3	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998439	0,998436	0,998437	0,999990
4	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,997198	0,997186	0,997187	0,999929
5	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998749	0,998749	0,998749	0,999992
Avg	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998315	0,998311	0,998312	0,999976

Table A.11. RF-Bitcoinfog (max_depth:4, min_samples_split:10, min_samples_leaf:2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996563	0,996563	0,996562	0,999938
2	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996258	0,996248	0,996249	0,999885
3	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996888	0,996873	0,996874	0,999913
4	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,995314	0,995311	0,995312	0,999752
5	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996565	0,996561	0,996562	0,999946
Avg	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996318	0,996311	0,996312	0,999887

Table A.12. RF-Bitcoinfog (max_depth:8, min_samples_split:10, min_samples_leaf:2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998749	0,998749	0,998749	0,999989
2	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,997818	0,997811	0,997812	0,999973
3	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998439	0,998436	0,998437	0,999987
4	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,997198	0,997186	0,997187	0,999924
5	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998749	0,998749	0,998749	0,999992
Avg	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998191	0,998186	0,998187	0,999973

Table A.13. RF-Bitcoinfog (max_depth:4, min_samples_split:2, min_samples_leaf:2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996563	0,996563	0,996562	0,999940
2	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996258	0,996248	0,996249	0,999885
3	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996888	0,996873	0,996874	0,999912
4	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,995314	0,995311	0,995312	0,999756
5	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996565	0,996561	0,996562	0,999945
Avg	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996318	0,996311	0,996312	0,999888

Table A.14. RF-Bitcoinfog (max_depth:8, min_samples_split:2, min_samples_leaf:4), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,999062	0,999062	0,999062	0,999991
2	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,997503	0,997499	0,997499	0,999973
3	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998439	0,998436	0,998437	0,999985
4	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,997198	0,997186	0,997187	0,999921
5	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998749	0,998749	0,998749	0,999992
Avg	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998190	0,998186	0,998187	0,999972

Table A.15. RF-Bitcoinfog (max_depth:4, min_samples_split:10, min_samples_leaf:4), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996563	0,996563	0,996562	0,999939
2	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996258	0,996248	0,996249	0,999885
3	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996888	0,996873	0,996874	0,999912
4	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,995314	0,995311	0,995312	0,999752
5	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996565	0,996561	0,996562	0,999944
Avg	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,996318	0,996311	0,996312	0,999886

Table A.16. RF-Bitcoinfog (max_depth:8, min_samples_split:10, min_samples_leaf:4), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,999062	0,999062	0,999062	0,999991
2	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,997503	0,997499	0,997499	0,999974
3	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998439	0,998436	0,998437	0,999985
4	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,997198	0,997186	0,997187	0,999909
5	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998749	0,998749	0,998749	0,999991
Avg	**Bitcoinfog**	$RF^{\text{EDF,EIF.}}$	0,998190	0,998186	0,998187	0,999970

Table A.17. RF-Bitcoinfog (max_depth:4, min_samples_split:2 min_samples_leaf:2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,943026	0,941890	0,941838	0,985026
2	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,947008	0,945331	0,945262	0,984474
3	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,946271	0,944084	0,943995	0,985496
4	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,936883	0,934087	0,933958	0,982917
5	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,936698	0,935347	0,935264	0,982046
Avg	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,941977	0,940148	0,940063	0,983992

Table A.18. RF-Bitcoinfog (max_depth:8, min_samples_split:2 min_samples_leaf:2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,956103	0,955943	0,955933	0,989630
2	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,962023	0,961572	0,961553	0,989033
3	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,961105	0,960943	0,960934	0,989044
4	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,954963	0,954386	0,954361	0,989111
5	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,955171	0,955012	0,954996	0,990048
Avg	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,957873	0,957571	0,957555	0,989373

Table A.19. RF-Bitcoinfog (max_depth:4, min_samples_split:10 min_samples_leaf:2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,943026	0,941890	0,941838	0,985003
2	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,947008	0,945331	0,945262	0,984532
3	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,946271	0,944084	0,943995	0,985470
4	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,936883	0,934087	0,933958	0,982917
5	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,936698	0,935347	0,935264	0,982010
Avg	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,941977	0,940148	0,940063	0,983986

Table A.20. RF-Bitcoinfog (max_depth:8, min_samples_split:10 min_samples_leaf:2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,957262	0,957191	0,957185	0,989548
2	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,960281	0,960007	0,959994	0,989166
3	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,962011	0,961880	0,961872	0,988916
4	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,955256	0,954698	0,954674	0,989159
5	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,954486	0,954385	0,954372	0,989884
Avg	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,957859	0,957632	0,957619	0,989335

Table A.21. RF-Bitcoinfog (max_depth:4, min_samples_split:2 min_samples_leaf:4), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,943026	0,941890	0,941838	0,985019
2	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,947008	0,945331	0,945262	0,984486
3	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,945996	0,943772	0,943681	0,985450
4	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,936883	0,934087	0,933958	0,982929
5	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,936416	0,935035	0,934950	0,981997
Avg	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,941866	0,940023	0,939938	0,983976

Table A.22. RF-Bitcoinfog (max_depth:8, min_samples_split:2 min_samples_leaf:4), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,956637	0,956566	0,956560	0,989460
2	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,961561	0,961258	0,961243	0,989265
3	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,962011	0,961880	0,961872	0,988904
4	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,954630	0,954073	0,954048	0,989100
5	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,952347	0,952199	0,952184	0,989862
Avg	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,957437	0,957195	0,957182	0,989318

Table A.23. RF-Bitcoinfog (max_depth:4, min_samples_split:10 min_samples_leaf:4), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,943026	0,941890	0,941838	0,985000
2	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,947008	0,945331	0,945262	0,984497
3	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,946271	0,944084	0,943995	0,985448
4	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,936883	0,934087	0,933958	0,982927
5	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,936416	0,935035	0,934950	0,981955
Avg	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,941921	0,940086	0,940001	0,983965

Table A.24. RF-Bitcoinfog (max_depth:8, min_samples_split:10 min_samples_leaf:4), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,955998	0,955941	0,955936	0,989512
2	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,960638	0,960320	0,960305	0,989174
3	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,961084	0,960943	0,960934	0,988917
4	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,954630	0,954073	0,954048	0,989087
5	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,952932	0,952822	0,952810	0,989831
Avg	**Bitcoinfog**	$RF^{\text{EDF.}}$	0,957056	0,956820	0,956807	0,989304

Table A.25. XGB-Bitcoinfog (max_depth:4, learning_rate:0.1), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999687	0,999687	0,999687	0,999998
2	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999687	0,999687	0,999687	0,999990
3	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999376	0,999374	0,999374	0,999965
4	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999064	0,999061	0,999062	0,999990
5	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999062	0,999062	0,999062	0,999997
Avg	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999375	0,999374	0,999374	0,999988

Table A.26. XGB-Bitcoinfog (max_depth:8, learning_rate:0.1), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,997193	0,997186	0,997187	0,999961
2	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,997513	0,997498	0,997499	0,999957
3	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,998125	0,998125	0,998124	0,999987
4	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,996562	0,996562	0,996562	0,999520
5	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,997187	0,997187	0,997187	0,999959
Avg	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,997316	0,997312	0,997312	0,999877

Table A.27. XGB-Bitcoinfog (max_depth:4, learning_rate:0.2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999374	0,999374	0,999374	0,999998
2	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999687	0,999687	0,999687	0,999990
3	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999376	0,999374	0,999374	0,999990
4	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999064	0,999061	0,999062	0,999990
5	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,998437	0,998437	0,998437	0,999996
Avg	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,999188	0,999187	0,999187	0,999993

Table A.28. XGB-Bitcoinfog (max_depth:8, learning_rate:0.2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,996878	0,996874	0,996874	0,999949
2	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,997508	0,997498	0,997499	0,999957
3	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,998125	0,998125	0,998124	0,999977
4	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,996562	0,996562	0,996562	0,999513
5	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,997187	0,997187	0,997187	0,999967
Avg	**Bitcoinfog**	$XGB^{\text{EIF.}}$	0,997252	0,997249	0,997249	0,999873

Table A.29. XGB-Bitcoinfog (max_depth:4, learning_rate:0.1), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999687	0,999687	0,999687	0,999998
2	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999687	0,999687	0,999687	0,999987
3	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999376	0,999374	0,999374	0,999985
4	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999376	0,999374	0,999374	0,999986
5	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,998749	0,998749	0,998749	0,999996
Avg	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999375	0,999374	0,999374	0,999990

Table A.30. XGB-Bitcoinfog (max_depth:8, learning_rate:0.1), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,997503	0,997499	0,997499	0,999967
2	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,997823	0,997811	0,997812	0,999967
3	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,998124	0,998124	0,998124	0,999985
4	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,996562	0,996562	0,996562	0,999010
5	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,996874	0,996874	0,996874	0,999962
Avg	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,997378	0,997374	0,997374	0,999778

Table A.31. XGB-Bitcoinfog (max_depth:4, learning_rate:0.2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999687	0,999687	0,999687	0,999999
2	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999687	0,999687	0,999687	0,999978
3	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999376	0,999374	0,999374	0,999982
4	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999376	0,999374	0,999374	0,999983
5	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,998749	0,998749	0,998749	0,999996
Avg	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,999375	0,999374	0,999374	0,999987

Table A.32. XGB-Bitcoinfog (max_depth:8, learning_rate:0.2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,996878	0,996874	0,996874	0,999962
2	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,997193	0,997186	0,997187	0,999948
3	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,998124	0,998124	0,998124	0,999983
4	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,996562	0,996562	0,996562	0,999485
5	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,997812	0,997812	0,997812	0,999966
Avg	**Bitcoinfog**	$XGB^{\text{EDF,EIF.}}$	0,997314	0,997312	0,997312	0,999869

Table A.33. XGB-Bitcoinfog (max_depth:4, learning_rate:0.1), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,962825	0,962814	0,962812	0,990135
2	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,960852	0,960632	0,960620	0,989810
3	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,962314	0,962192	0,962185	0,990031
4	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,959628	0,959382	0,959369	0,988171
5	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,958815	0,958757	0,958748	0,991372
Avg	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,960887	0,960755	0,960747	0,989904

Table A.34. XGB-Bitcoinfog (max_depth:8, learning_rate:0.1), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,959082	0,959064	0,959062	0,989528
2	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,962378	0,962193	0,962183	0,988260
3	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,958793	0,958753	0,958749	0,989004
4	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,955472	0,955010	0,954989	0,987457
5	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,958237	0,958135	0,958123	0,990187
Avg	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,958792	0,958631	0,958621	0,988887

Table A.35. XGB-Bitcoinfog (max_depth:4, learning_rate:0.2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,959398	0,959377	0,959374	0,989641
2	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,959813	0,959692	0,959685	0,987543
3	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,957262	0,957191	0,957185	0,987668
4	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,956751	0,956568	0,956558	0,987369
5	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,957579	0,957508	0,957498	0,989565
Avg	**Bitcoinfog**	$XGB^{\text{EDF.}}$	0,958161	0,958067	0,958060	0,988357

Table A.36. RF-Helix (max_depth:4, min_samples_split:2, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EIF.}}$	0,999373	0,999838	0,999605	0,999999
3	**Helix**	$RF^{\text{EIF.}}$	0,999210	0,999210	0,999210	0,999999
4	**Helix**	$RF^{\text{EIF.}}$	0,998666	0,999363	0,999014	0,999998
4	**Helix**	$RF^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
Avg	**Helix**	$RF^{\text{EIF.}}$	0,999292	0,999524	0,999408	0,999999

Table A.37. RF-Helix (max_depth:8, min_samples_split:2, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	1
2	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
3	**Helix**	$RF^{\text{EIF.}}$	0,999443	0,998977	0,999210	0,999999
4	**Helix**	$RF^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999998
5	**Helix**	$RF^{\text{EIF.}}$	0,999838	0,999372	0,999605	0,999999
Avg	**Helix**	$RF^{\text{EIF.}}$	0,999589	0,999542	0,999565	0,999999

Table A.38. RF-Helix (max_depth:4, min_samples_split:10, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EIF.}}$	0,999373	0,999838	0,999605	0,999999
3	**Helix**	$RF^{\text{EIF.}}$	0,999210	0,999210	0,999210	0,999999
4	**Helix**	$RF^{\text{EIF.}}$	0,998666	0,999363	0,999014	0,999998
5	**Helix**	$RF^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
Avg	**Helix**	$RF^{\text{EIF.}}$	0,999292	0,999524	0,999408	0,999999

Table A.39. RF-Helix (max_depth:8, min_samples_split:10, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
3	**Helix**	$RF^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
4	**Helix**	$RF^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999998
5	**Helix**	$RF^{\text{EIF.}}$	0,999838	0,999372	0,999605	0,999999
Avg	**Helix**	$RF^{\text{EIF.}}$	0,999605	0,999605	0,999605	0,999999

Table A.40. RF-Helix (max_depth:4, min_samples_split:2, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EIF.}}$	0,999060	0,999758	0,999408	0,999999
3	**Helix**	$RF^{\text{EIF.}}$	0,999210	0,999210	0,999210	0,999999
4	**Helix**	$RF^{\text{EIF.}}$	0,998666	0,999363	0,999014	0,999998
5	**Helix**	$RF^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
Avg	**Helix**	$RF^{\text{EIF.}}$	0,999229	0,999508	0,999368	0,999999

Table A.41. RF-Helix (max_depth:8, min_samples_split:2, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
3	**Helix**	$RF^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
4	**Helix**	$RF^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999998
5	**Helix**	$RF^{\text{EIF.}}$	0,999838	0,999372	0,999605	1
Avg	**Helix**	$RF^{\text{EIF.}}$	0,999605	0,999605	0,999605	0,999999

Table A.42. RF-Helix (max_depth:4, min_samples_split:10, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EIF.}}$	0,999060	0,999758	0,999408	0,999999
3	**Helix**	$RF^{\text{EIF.}}$	0,999210	0,999210	0,999210	0,999999
4	**Helix**	$RF^{\text{EIF.}}$	0,998666	0,999363	0,999014	0,999998
5	**Helix**	$RF^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
Avg	**Helix**	$RF^{\text{EIF.}}$	0,999229	0,999508	0,999368	0,999999

Table A.43. RF-Helix (max_depth:8, min_samples_split:10, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EIF.}}$	0,999686	0,999919	0,999802	0,999999
3	**Helix**	$RF^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
4	**Helix**	$RF^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999998
5	**Helix**	$RF^{\text{EIF.}}$	0,999838	0,999372	0,999605	1
Avg	**Helix**	$RF^{\text{EDF.}}$	0,999605	0,999605	0,999605	0,999999

Table A.44. XGB-Helix (max_depth:4, learning_rate:0.1), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EIF.}}$	1	1	1	1
2	**Helix**	$XGB^{\text{EIF.}}$	0,999838	0,999372	0,999605	1
3	**Helix**	$XGB^{\text{EIF.}}$	0,999686	0,999919	0,999802	1
4	**Helix**	$XGB^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999999
5	**Helix**	$XGB^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
Avg	**Helix**	$XGB^{\text{EIF.}}$	0,999668	0,999621	0,999644	0,999999

Table A.45. XGB-Helix (max_depth:8, learning_rate:0.1), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999992
2	**Helix**	$XGB^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
3	**Helix**	$XGB^{\text{EIF.}}$	0,999060	0,999758	0,999408	0,999996
4	**Helix**	$XGB^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999989
5	**Helix**	$XGB^{\text{EIF.}}$	0,998340	0,998108	0,998224	0,999994
Avg	**Helix**	$XGB^{\text{EIF.}}$	0,999101	0,999241	0,999171	0,999994

Table A.46. XGB-Helix (max_depth:4, learning_rate:0.2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EIF.}}$	1	1	1	1
2	**Helix**	$XGB^{\text{EIF.}}$	0,999838	0,999372	0,999605	1
3	**Helix**	$XGB^{\text{EIF.}}$	0,999686	0,999919	0,999802	1
4	**Helix**	$XGB^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999999
5	**Helix**	$XGB^{\text{EIF.}}$	0,999919	0,999686	0,999802	0,999999
Avg	**Helix**	$XGB^{\text{EIF.}}$	0,999747	0,999700	0,999723	0,999999

Table A.47. XGB-Helix (max_depth:8, learning_rate:0.2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999994
2	**Helix**	$XGB^{\text{EIF.}}$	0,999524	0,999291	0,999408	0,999999
3	**Helix**	$XGB^{\text{EIF.}}$	0,999060	0,999758	0,999408	0,999993
4	**Helix**	$XGB^{\text{EIF.}}$	0,999291	0,999524	0,999408	0,999994
5	**Helix**	$XGB^{\text{EIF.}}$	0,998340	0,998108	0,998224	0,999994
Avg	**Helix**	$XGB^{\text{EIF.}}$	0,999101	0,999241	0,999171	0,999995

Table A.48. RF-Helix (max_depth:4, min_samples_split:2, min_samples_leaf:2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999373	0,999838	0,999605	0,999999
3	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999210	0,999210	0,999210	0,999999
4	**Helix**	$RF^{\text{EDF,EIF.}}$	0,998666	0,999363	0,999014	0,999998
5	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999605	0,999605	0,999605	0,999999
Avg	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999308	0,999587	0,999447	0,999999

Table A.49. RF-Helix (max_depth:8, min_samples_split:2, min_samples_leaf:2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	1
2	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
3	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999524	0,999291	0,999408	0,999999
4	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999291	0,999524	0,999408	0,999998
5	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999838	0,999372	0,999605	0,999999
Avg	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999605	0,999605	0,999605	0,999999

Table A.50. RF-Helix (max_depth:4, min_samples_split:10, min_samples_leaf:2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999373	0,999838	0,999605	0,999999
3	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999210	0,999210	0,999210	0,999999
4	**Helix**	$RF^{\text{EDF,EIF.}}$	0,998666	0,999363	0,999014	0,999998
5	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999605	0,999605	0,999605	0,999999
Avg	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999308	0,999587	0,999447	0,999999

Table A.51. RF-Helix (max_depth:8, min_samples_split:10, min_samples_leaf:2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
3	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999524	0,999291	0,999408	0,999999
4	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999291	0,999524	0,999408	0,999998
5	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999838	0,999372	0,999605	0,999999
Avg	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999605	0,999605	0,999605	0,999999

Table A.52. RF-Helix (max_depth:4, min_samples_split:2, min_samples_leaf:4), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999373	0,999838	0,999605	0,999999
3	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999210	0,999210	0,999210	0,999999
4	**Helix**	$RF^{\text{EDF,EIF.}}$	0,998353	0,999282	0,998817	0,999998
5	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999605	0,999605	0,999605	0,999999
Avg	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999245	0,999571	0,999408	0,999999

Table A.53. RF-Helix (max_depth:8, min_samples_split:2, min_samples_leaf:4), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
3	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999524	0,999291	0,999408	0,999999
4	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999291	0,999524	0,999408	0,999998
5	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999838	0,999372	0,999605	1
Avg	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999605	0,999605	0,999605	0,999999

Table A.54. RF-Helix (max_depth:4, min_samples_split:10, min_samples_leaf:4), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999060	0,999758	0,999408	0,999999
3	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999210	0,999210	0,999210	0,999999
4	**Helix**	$RF^{\text{EDF,EIF.}}$	0,998353	0,999282	0,998817	0,999998
5	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
Avg	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999199	0,999618	0,999408	0,999999

Table A.55. RF-Helix (max_depth:4, min_samples_split:8, min_samples_leaf:4), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
2	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	0,999999
3	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999524	0,999291	0,999408	0,999999
4	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999291	0,999524	0,999408	0,999998
5	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999838	0,999372	0,999605	1
Avg	**Helix**	$RF^{\text{EDF,EIF.}}$	0,999605	0,999605	0,999605	0,999999

Table A.56. XGB-Helix (max_depth:4, learning_rate:0.1), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EDF,EIF.}}$	1	1	1	1
2	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999838	0,999372	0,999605	1
3	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	1
4	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999291	0,999524	0,999408	0,999999
5	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999605	0,999605	0,999605	0,999999
Avg	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999684	0,999684	0,999684	0,999999

Table A.57. XGB-Helix (max_depth:8, learning_rate:0.1), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999291	0,999524	0,999408	0,999992
2	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999524	0,999291	0,999408	0,999999
3	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999060	0,999758	0,999408	0,999996
4	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999291	0,999524	0,999408	0,999989
5	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,998340	0,998108	0,998224	0,999994
Avg	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999101	0,999241	0,999171	0,999994

Table A.58. XGB-Helix (max_depth:4, learning_rate:0.2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EDF,EIF.}}$	1	1	1	1
2	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999919	0,999686	0,999802	1
3	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999686	0,999919	0,999802	1
4	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,998978	0,999444	0,999211	0,999999
5	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999919	0,999686	0,999802	1
Avg	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999700	0,999747	0,999723	0,999999

Table A.59. XGB-Helix (max_depth:8, learning_rate:0.2), EDF,EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999291	0,999524	0,999408	0,999994
2	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999524	0,999291	0,999408	0,999999
3	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999060	0,999758	0,999408	0,999993
4	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,998978	0,999444	0,999211	0,999991
5	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,998340	0,998108	0,998224	0,999994
Avg	**Helix**	$XGB^{\text{EDF,EIF.}}$	0,999039	0,999225	0,999132	0,999994

Table A.60. RF-Helix (max_depth:4, min_samples_split:2, min_samples_leaf:2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF.}}$	0,922479	0,958558	0,938948	0,990624
2	**Helix**	$RF^{\text{EDF.}}$	0,913240	0,954319	0,931680	0,991289
3	**Helix**	$RF^{\text{EDF.}}$	0,912420	0,956185	0,931909	0,992389
4	**Helix**	$RF^{\text{EDF.}}$	0,915780	0,960592	0,935702	0,992367
5	**Helix**	$RF^{\text{EDF.}}$	0,911470	0,952889	0,930032	0,991748
Avg	**Helix**	$RF^{\text{EDF.}}$	0,915078	0,956509	0,933654	0,991683

Table A.61. RF-Helix (max_depth:8, min_samples_split:2, min_samples_leaf:2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF.}}$	0,944719	0,966202	0,954942	0,994604
2	**Helix**	$RF^{\text{EDF.}}$	0,945104	0,969960	0,956838	0,994709
3	**Helix**	$RF^{\text{EDF.}}$	0,941540	0,969692	0,954715	0,995503
4	**Helix**	$RF^{\text{EDF.}}$	0,945437	0,970899	0,957440	0,995030
5	**Helix**	$RF^{\text{EDF.}}$	0,943347	0,968299	0,955120	0,995443
Avg	**Helix**	$RF^{\text{EDF.}}$	0,944029	0,969010	0,955811	0,995058

Table A.62. RF-Helix (max_depth:4, min_samples_split:10, min_samples_leaf:2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF.}}$	0,921039	0,958074	0,937894	0,990631
2	**Helix**	$RF^{\text{EDF.}}$	0,913240	0,954319	0,931680	0,991279
3	**Helix**	$RF^{\text{EDF.}}$	0,911963	0,956023	0,931564	0,992374
4	**Helix**	$RF^{\text{EDF.}}$	0,915780	0,960592	0,935702	0,992369
5	**Helix**	$RF^{\text{EDF.}}$	0,911470	0,952889	0,930032	0,991777
Avg	**Helix**	$RF^{\text{EDF.}}$	0,914698	0,956380	0,933374	0,991686

Table A.63. RF-Helix (max_depth:8, min_samples_split:10, min_samples_leaf:2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF.}}$	0,944830	0,966516	0,955144	0,994587
2	**Helix**	$RF^{\text{EDF.}}$	0,946157	0,970282	0,957568	0,994821
3	**Helix**	$RF^{\text{EDF.}}$	0,941026	0,969530	0,954353	0,995531
4	**Helix**	$RF^{\text{EDF.}}$	0,945437	0,970899	0,957440	0,995046
5	**Helix**	$RF^{\text{EDF.}}$	0,943459	0,968613	0,955321	0,995411
Avg	**Helix**	$RF^{\text{EDF.}}$	0,944181	0,969168	0,955965	0,995079

Table A.64. RF-Helix (max_depth:4, min_samples_split:10, min_samples_leaf:2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF.}}$	0,944830	0,966516	0,955144	0,994587
2	**Helix**	$RF^{\text{EDF.}}$	0,946157	0,970282	0,957568	0,994821
3	**Helix**	$RF^{\text{EDF.}}$	0,941026	0,969530	0,954353	0,995531
4	**Helix**	$RF^{\text{EDF.}}$	0,945437	0,970899	0,957440	0,995046
5	**Helix**	$RF^{\text{EDF.}}$	0,943459	0,968613	0,955321	0,995411
Avg	**Helix**	$RF^{\text{EDF.}}$	0,944181	0,969168	0,955965	0,995079

Table A.65. RF-Helix (max_depth:8, min_samples_split:2, min_samples_leaf:4), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF.}}$	0,945095	0,966596	0,955327	0,994503
2	**Helix**	$RF^{\text{EDF.}}$	0,946046	0,969968	0,957366	0,994779
3	**Helix**	$RF^{\text{EDF.}}$	0,941685	0,969458	0,954696	0,995425
4	**Helix**	$RF^{\text{EDF.}}$	0,945699	0,970980	0,957622	0,994980
5	**Helix**	$RF^{\text{EDF.}}$	0,943720	0,968693	0,955503	0,995386
Avg	**Helix**	$RF^{\text{EDF.}}$	0,944449	0,969139	0,956103	0,995015

Table A.66. RF-Helix (max_depth:4, min_samples_split:10, min_samples_leaf:4), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF.}}$	0,921039	0,958074	0,937894	0,990626
2	**Helix**	$RF^{\text{EDF.}}$	0,913240	0,954319	0,931680	0,991360
3	**Helix**	$RF^{\text{EDF.}}$	0,911963	0,956023	0,931564	0,992370
4	**Helix**	$RF^{\text{EDF.}}$	0,915780	0,960592	0,935702	0,992377
5	**Helix**	$RF^{\text{EDF.}}$	0,911470	0,952889	0,930032	0,991770
Avg	**Helix**	$RF^{\text{EDF.}}$	0,914698	0,956380	0,933374	0,991701

Table A.67. RF-Helix (max_depth:8, min_samples_split:10, min_samples_leaf:4), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$RF^{\text{EDF.}}$	0,944830	0,966516	0,955144	0,994540
2	**Helix**	$RF^{\text{EDF.}}$	0,945782	0,969888	0,957184	0,994762
3	**Helix**	$RF^{\text{EDF.}}$	0,941396	0,969925	0,954735	0,995457
4	**Helix**	$RF^{\text{EDF.}}$	0,945548	0,971213	0,957641	0,994993
5	**Helix**	$RF^{\text{EDF.}}$	0,943049	0,968765	0,955159	0,995379
Avg	**Helix**	$RF^{\text{EDF.}}$	0,944121	0,969261	0,955973	0,995026

Table A.68. XGB-Helix (max_depth:4, learning_rate:0.1), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EDF.}}$	0,953621	0,963464	0,958437	0,990650
2	**Helix**	$XGB^{\text{EDF.}}$	0,953743	0,970585	0,961854	0,992115
3	**Helix**	$XGB^{\text{EDF.}}$	0,954609	0,971688	0,962830	0,993460
4	**Helix**	$XGB^{\text{EDF.}}$	0,957269	0,972179	0,964483	0,993574
5	**Helix**	$XGB^{\text{EDF.}}$	0,953487	0,974180	0,963364	0,995503
Avg	**Helix**	$XGB^{\text{EDF.}}$	0,954546	0,970419	0,962194	0,993060

Table A.69. XGB-Helix (max_depth:8, learning_rate:0.1), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EDF.}}$	0,960684	0,966027	0,963325	0,991265
2	**Helix**	$XGB^{\text{EDF.}}$	0,958970	0,970395	0,964542	0,994951
3	**Helix**	$XGB^{\text{EDF.}}$	0,956202	0,968181	0,962036	0,994282
4	**Helix**	$XGB^{\text{EDF.}}$	0,960570	0,970564	0,965460	0,994662
5	**Helix**	$XGB^{\text{EDF.}}$	0,961645	0,972295	0,966849	0,995264
Avg	**Helix**	$XGB^{\text{EDF.}}$	0,959614	0,969493	0,964442	0,994085

Table A.70. XGB-Helix (max_depth:4, learning_rate:0.2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EDF.}}$	0,955121	0,966450	0,960646	0,990330
2	**Helix**	$XGB^{\text{EDF.}}$	0,957590	0,970853	0,964031	0,993315
3	**Helix**	$XGB^{\text{EDF.}}$	0,953516	0,967688	0,960382	0,991448
4	**Helix**	$XGB^{\text{EDF.}}$	0,958665	0,972582	0,965415	0,992031
5	**Helix**	$XGB^{\text{EDF.}}$	0,954633	0,970278	0,962189	0,989353
Avg	**Helix**	$XGB^{\text{EDF.}}$	0,955905	0,969570	0,962533	0,991295

Table A.71. XGB-Helix (max_depth:8, learning_rate:0.2), EDF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Helix**	$XGB^{\text{EDF.}}$	0,960858	0,964072	0,962454	0,991293
2	**Helix**	$XGB^{\text{EDF.}}$	0,960637	0,969157	0,964820	0,993419
3	**Helix**	$XGB^{\text{EDF.}}$	0,955165	0,965042	0,959998	0,988025
4	**Helix**	$XGB^{\text{EDF.}}$	0,962173	0,969873	0,965960	0,987730
5	**Helix**	$XGB^{\text{EDF.}}$	0,959481	0,971963	0,965554	0,992283
Avg	**Helix**	$XGB^{\text{EDF.}}$	0,959663	0,968022	0,963757	0,990550

Table A.72. RF-Bitcoin (max_depth:4, min_samples_split:2, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$RF^{\text{EIF.}}$	0,987974	0,992033	0,989945	0,999674
2	**Bitcoin**	$RF^{\text{EIF.}}$	0,991699	0,993413	0,992546	0,999794
3	**Bitcoin**	$RF^{\text{EIF.}}$	0,992574	0,994291	0,993423	0,999874
4	**Bitcoin**	$RF^{\text{EIF.}}$	0,988944	0,991845	0,990365	0,999807
5	**Bitcoin**	$RF^{\text{EIF.}}$	0,991531	0,991778	0,991654	0,999778
Avg	**Bitcoin**	$RF^{\text{EIF.}}$	0,990544	0,992672	0,991587	0,999785

Table A.73. RF-Bitcoin (max_depth:8, min_samples_split:2, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$RF^{\text{EIF.}}$	0,995205	0,996926	0,996055	0,999933
2	**Bitcoin**	$RF^{\text{EIF.}}$	0,996873	0,997867	0,997367	0,999979
3	**Bitcoin**	$RF^{\text{EIF.}}$	0,997995	0,998494	0,998243	0,999990
4	**Bitcoin**	$RF^{\text{EIF.}}$	0,998310	0,999059	0,998683	0,999975
5	**Bitcoin**	$RF^{\text{EIF.}}$	0,998242	0,998242	0,998242	0,999963
Avg	**Bitcoin**	$RF^{\text{EIF.}}$	0,997325	0,998118	0,997718	0,999968

Table A.74. RF-Bitcoin (max_depth:4, min_samples_split:10, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$RF^{\text{EIF.}}$	0,987433	0,991719	0,989510	0,999667
2	**Bitcoin**	$RF^{\text{EIF.}}$	0,991145	0,993099	0,992109	0,999797
3	**Bitcoin**	$RF^{\text{EIF.}}$	0,992574	0,994291	0,993423	0,999874
4	**Bitcoin**	$RF^{\text{EIF.}}$	0,988944	0,991845	0,990365	0,999804
5	**Bitcoin**	$RF^{\text{EIF.}}$	0,991213	0,991213	0,991213	0,999773
Avg	**Bitcoin**	$RF^{\text{EIF.}}$	0,990262	0,992434	0,991324	0,999783

Table A.75. RF-Bitcoin (max_depth:8, min_samples_split:10, min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$RF^{\text{EIF.}}$	0,995525	0,997490	0,996494	0,999939
2	**Bitcoin**	$RF^{\text{EIF.}}$	0,996873	0,997867	0,997367	0,999975
3	**Bitcoin**	$RF^{\text{EIF.}}$	0,997433	0,998181	0,997805	0,999991
4	**Bitcoin**	$RF^{\text{EIF.}}$	0,998310	0,999059	0,998683	0,999977
5	**Bitcoin**	$RF^{\text{EIF.}}$	0,997679	0,997929	0,997803	0,999957
Avg	**Bitcoin**	$RF^{\text{EIF.}}$	0,997164	0,998105	0,997631	0,999968

Table A.76. RF-Bitcoin (max_depth:4, min_samples_split:2, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$RF^{\text{EIF.}}$	0,987433	0,991719	0,989510	0,999667
2	**Bitcoin**	$RF^{\text{EIF.}}$	0,991699	0,993413	0,992546	0,999795
3	**Bitcoin**	$RF^{\text{EIF.}}$	0,992574	0,994291	0,993423	0,999873
4	**Bitcoin**	$RF^{\text{EIF.}}$	0,988944	0,991845	0,990365	0,999802
5	**Bitcoin**	$RF^{\text{EIF.}}$	0,990895	0,990648	0,990771	0,999766
Avg	**Bitcoin**	$RF^{\text{EIF.}}$	0,990309	0,992383	0,991323	0,999781

Table A.77. RF-Bitcoin (max_depth:8, min_samples_split:2, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$RF^{\text{EIF.}}$	0,994885	0,996361	0,995616	0,999935
2	**Bitcoin**	$RF^{\text{EIF.}}$	0,996873	0,997867	0,997367	0,999969
3	**Bitcoin**	$RF^{\text{EIF.}}$	0,997433	0,998181	0,997805	0,999991
4	**Bitcoin**	$RF^{\text{EIF.}}$	0,997191	0,998432	0,997806	0,999970
5	**Bitcoin**	$RF^{\text{EIF.}}$	0,997364	0,997364	0,997364	0,999959
Avg	**Bitcoin**	$RF^{\text{EIF.}}$	0,996749	0,997641	0,997191	0,999965

Table A.78. RF-Bitcoin (max_depth:4, min_samples_split:10, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$RF^{\text{EIF.}}$	0,987433	0,991719	0,989510	0,999665
2	**Bitcoin**	$RF^{\text{EIF.}}$	0,991145	0,993099	0,992109	0,999794
3	**Bitcoin**	$RF^{\text{EIF.}}$	0,992574	0,994291	0,993423	0,999873
4	**Bitcoin**	$RF^{\text{EIF.}}$	0,988944	0,991845	0,990365	0,999803
5	**Bitcoin**	$RF^{\text{EIF.}}$	0,991213	0,991213	0,991213	0,999767
Avg	**Bitcoin**	$RF^{\text{EIF.}}$	0,990262	0,992434	0,991324	0,999780

Table A.79. RF-Bitcoin (max_depth:8, min_samples_split:10, min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$RF^{\text{EIF.}}$	0,994566	0,995797	0,995176	0,999932
2	**Bitcoin**	$RF^{\text{EIF.}}$	0,996873	0,997867	0,997367	0,999967
3	**Bitcoin**	$RF^{\text{EIF.}}$	0,997433	0,998181	0,997805	0,999990
4	**Bitcoin**	$RF^{\text{EIF.}}$	0,997191	0,998432	0,997806	0,999972
5	**Bitcoin**	$RF^{\text{EIF.}}$	0,996239	0,996736	0,996487	0,999955
Avg	**Bitcoin**	$RF^{\text{EIF.}}$	0,996460	0,997403	0,996928	0,999963

Table A.80. XGB-Helix (max_depth:4, learning_rate:0.1), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$XGB^{\text{EIF.}}$	0,997929	0,997679	0,997804	0,999976
2	**Bitcoin**	$XGB^{\text{EIF.}}$	0,999121	0,999121	0,999121	0,999998
3	**Bitcoin**	$XGB^{\text{EIF.}}$	0,999435	0,999686	0,999560	1
4	**Bitcoin**	$XGB^{\text{EIF.}}$	1	1	1	1
5	**Bitcoin**	$XGB^{\text{EIF.}}$	0,998557	0,998807	0,998682	0,999980
Avg	**Bitcoin**	$XGB^{\text{EIF.}}$	0,999008	0,999059	0,999033	0,999991

Table A.81. XGB-Helix (max_depth:8, learning_rate:0.1), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$XGB^{\text{EIF.}}$	0,993929	0,994668	0,994297	0,999423
2	**Bitcoin**	$XGB^{\text{EIF.}}$	0,997615	0,997113	0,997363	0,999897
3	**Bitcoin**	$XGB^{\text{EIF.}}$	0,996363	0,994853	0,995601	0,999843
4	**Bitcoin**	$XGB^{\text{EIF.}}$	0,994858	0,992844	0,993838	0,999294
5	**Bitcoin**	$XGB^{\text{EIF.}}$	0,994096	0,993598	0,993846	0,999757
Avg	**Bitcoin**	$XGB^{\text{EIF.}}$	0,995372	0,994615	0,994989	0,999643

Table A.82. XGB-Helix (max_depth:4, learning_rate:0.2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$XGB^{\text{EIF.}}$	0,997929	0,997679	0,997804	0,999973
2	**Bitcoin**	$XGB^{\text{EIF.}}$	0,999121	0,999121	0,999121	0,999997
3	**Bitcoin**	$XGB^{\text{EIF.}}$	1	1	1	1
4	**Bitcoin**	$XGB^{\text{EIF.}}$	1	1	1	1
5	**Bitcoin**	$XGB^{\text{EIF.}}$	0,998557	0,998807	0,998682	0,999979
Avg	**Bitcoin**	$XGB^{\text{EIF.}}$	0,999121	0,999121	0,999121	0,999990

Table A.83. XGB-Helix (max_depth:8, learning_rate:0.2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Bitcoin**	$XGB^{\text{EIF.}}$	0,993929	0,994668	0,994297	0,999623
2	**Bitcoin**	$XGB^{\text{EIF.}}$	0,996421	0,995669	0,996043	0,999902
3	**Bitcoin**	$XGB^{\text{EIF.}}$	0,996363	0,994853	0,995601	0,999959
4	**Bitcoin**	$XGB^{\text{EIF.}}$	0,994858	0,992844	0,993838	0,999439
5	**Bitcoin**	$XGB^{\text{EIF.}}$	0,993780	0,993033	0,993405	0,998800
Avg	**Bitcoin**	$XGB^{\text{EIF.}}$	0,995070	0,994213	0,994637	0,999544

Table A.84. RF-Ethereum (max_depth:4, min_samples_split:2 min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$RF^{\text{EIF.}}$	0.841173	0.848004	0.832735	0.937016
2	**Ethereum**	$RF^{\text{EIF.}}$	0.848785	0.857099	0.843782	0.942910
3	**Ethereum**	$RF^{\text{EIF.}}$	0.841869	0.850345	0.842781	0.935838
4	**Ethereum**	$RF^{\text{EIF.}}$	0.853904	0.862824	0.854124	0.945435
5	**Ethereum**	$RF^{\text{EIF.}}$	0.848554	0.857142	0.845232	0.944181
Avg	**Ethereum**	$RF^{\text{EIF.}}$	0.846857	0.855083	0.843731	0.941076

Table A.85. RF-Ethereum (max_depth:8, min_samples_split:2 min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$RF^{\text{EIF.}}$	0.867447	0.876069	0.869099	0.950827
2	**Ethereum**	$RF^{\text{EIF.}}$	0.880465	0.889238	0.882331	0.957661
3	**Ethereum**	$RF^{\text{EIF.}}$	0.854837	0.860220	0.856817	0.950596
4	**Ethereum**	$RF^{\text{EIF.}}$	0.871925	0.875202	0.873355	0.957857
5	**Ethereum**	$RF^{\text{EIF.}}$	0.883214	0.889813	0.885455	0.962444
Avg	**Ethereum**	$RF^{\text{EIF.}}$	0.871578	0.878108	0.873411	0.955877

Table A.86. RF-Ethereum (max_depth:4, min_samples_split:10 min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$RF^{\text{EIF.}}$	0.841173	0.848004	0.832735	0.936698
2	**Ethereum**	$RF^{\text{EIF.}}$	0.848785	0.857099	0.843782	0.942809
3	**Ethereum**	$RF^{\text{EIF.}}$	0.839787	0.848113	0.840827	0.935665
4	**Ethereum**	$RF^{\text{EIF.}}$	0.853904	0.862824	0.854124	0.945377
5	**Ethereum**	$RF^{\text{EIF.}}$	0.848554	0.857142	0.845232	0.944152
Avg	**Ethereum**	$RF^{\text{EIF.}}$	0.846440	0.854636	0.843340	0.940940

Table A.87. RF-Ethereum (max_depth:8, min_samples_split:10 min_samples_leaf:2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$RF^{\text{EIF.}}$	0.863685	0.872218	0.865304	0.949989
2	**Ethereum**	$RF^{\text{EIF.}}$	0.878733	0.887619	0.880481	0.957025
3	**Ethereum**	$RF^{\text{EIF.}}$	0.852972	0.858602	0.854980	0.949830
4	**Ethereum**	$RF^{\text{EIF.}}$	0.873778	0.877435	0.875342	0.957697
5	**Ethereum**	$RF^{\text{EIF.}}$	0.883214	0.889813	0.885455	0.961444
Avg	**Ethereum**	$RF^{\text{EIF.}}$	0.870476	0.877137	0.872312	0.955197

Table A.88. RF-Ethereum (max_depth:4, min_samples_split:2 min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$RF^{\text{EIF.}}$	0.842565	0.849622	0.834588	0.936669
2	**Ethereum**	$RF^{\text{EIF.}}$	0.848785	0.857099	0.843782	0.942607
3	**Ethereum**	$RF^{\text{EIF.}}$	0.841869	0.850345	0.842781	0.935390
4	**Ethereum**	$RF^{\text{EIF.}}$	0.856036	0.865056	0.856066	0.945189
5	**Ethereum**	$RF^{\text{EIF.}}$	0.848554	0.857142	0.845232	0.944123
Avg	**Ethereum**	$RF^{\text{EIF.}}$	0.847562	0.855853	0.844490	0.940796

Table A.89. RF-Ethereum (max_depth:8, min_samples_split:2 min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$RF^{\text{EIF.}}$	0.863685	0.872218	0.865304	0.949108
2	**Ethereum**	$RF^{\text{EIF.}}$	0.874678	0.883155	0.876587	0.956491
3	**Ethereum**	$RF^{\text{EIF.}}$	0.851058	0.854527	0.852536	0.948964
4	**Ethereum**	$RF^{\text{EIF.}}$	0.873778	0.877435	0.875342	0.957625
5	**Ethereum**	$RF^{\text{EIF.}}$	0.877700	0.884943	0.879894	0.962299
Avg	**Ethereum**	$RF^{\text{EIF.}}$	0.868180	0.874456	0.869933	0.954897

Table A.90. RF-Ethereum (max_depth:8, min_samples_split:2 min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$RF^{\text{EIF.}}$	0.842565	0.849622	0.834588	0.936582
2	**Ethereum**	$RF^{\text{EIF.}}$	0.848785	0.857099	0.843782	0.942621
3	**Ethereum**	$RF^{\text{EIF.}}$	0.837713	0.845881	0.838871	0.935376
4	**Ethereum**	$RF^{\text{EIF.}}$	0.853904	0.862824	0.854124	0.945247
5	**Ethereum**	$RF^{\text{EIF.}}$	0.848554	0.857142	0.845232	0.944152
Avg	**Ethereum**	$RF^{\text{EIF.}}$	0.846304	0.854514	0.843320	0.940796

Table A.91. RF-Ethereum (max_depth:8, min_samples_split:10 min_samples_leaf:4), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$RF^{\text{EIF.}}$	0.871209	0.879919	0.872893	0.949816
2	**Ethereum**	$RF^{\text{EIF.}}$	0.880465	0.889238	0.882331	0.956679
3	**Ethereum**	$RF^{\text{EIF.}}$	0.852935	0.856759	0.854530	0.948956
4	**Ethereum**	$RF^{\text{EIF.}}$	0.871823	0.875811	0.873492	0.957147
5	**Ethereum**	$RF^{\text{EIF.}}$	0.879669	0.887175	0.881854	0.961647
Avg	**Ethereum**	$RF^{\text{EIF.}}$	0.871220	0.877780	0.873020	0.954849

Table A.92. XGB-Ethereum (max_depth:4, learning_rate:0.1), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$XGB^{\text{EIF.}}$	0.873846	0.878243	0.875651	0.952626
2	**Ethereum**	$XGB^{\text{EIF.}}$	0.876058	0.878633	0.877227	0.951528
3	**Ethereum**	$XGB^{\text{EIF.}}$	0.877930	0.874949	0.876319	0.947613
4	**Ethereum**	$XGB^{\text{EIF.}}$	0.900684	0.903814	0.902084	0.967757
5	**Ethereum**	$XGB^{\text{EIF.}}$	0.905026	0.906452	0.905708	0.961662
Avg	**Ethereum**	$XGB^{\text{EIF.}}$	0.886709	0.888418	0.887398	0.956237

Table A.93. XGB-Ethereum (max_depth:8, learning_rate:0.1), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$XGB^{\text{EIF.}}$	0.872941	0.872941	0.872941	0.948884
2	**Ethereum**	$XGB^{\text{EIF.}}$	0.885245	0.883263	0.884200	0.952034
3	**Ethereum**	$XGB^{\text{EIF.}}$	0.878922	0.873721	0.875985	0.951145
4	**Ethereum**	$XGB^{\text{EIF.}}$	0.876337	0.877232	0.876771	0.958241
5	**Ethereum**	$XGB^{\text{EIF.}}$	0.901874	0.901379	0.901624	0.956690
Avg	**Ethereum**	$XGB^{\text{EIF.}}$	0.883064	0.881707	0.882304	0.953399

Table A.94. XGB-Ethereum (max_depth:4, learning_rate:0.2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$XGB^{\text{EIF.}}$	0.874064	0.877015	0.875379	0.950163
2	**Ethereum**	$XGB^{\text{EIF.}}$	0.879718	0.883097	0.881197	0.953161
3	**Ethereum**	$XGB^{\text{EIF.}}$	0.876217	0.872717	0.874303	0.943315
4	**Ethereum**	$XGB^{\text{EIF.}}$	0.894711	0.898944	0.896493	0.959698
5	**Ethereum**	$XGB^{\text{EIF.}}$	0.895900	0.895900	0.895900	0.959393
Avg	**Ethereum**	$XGB^{\text{EIF.}}$	0.884122	0.885535	0.884655	0.953146

Table A.95. XGB-Ethereum (max_depth:8, learning_rate:0.2), EIF.

Fold	Dataset	ML Model	Precision	Recall	F1 score	AUC
1	**Ethereum**	$XGB^{\text{EIF.}}$	0.877125	0.876177	0.876638	0.946349
2	**Ethereum**	$XGB^{\text{EIF.}}$	0.882735	0.882259	0.882494	0.953348
3	**Ethereum**	$XGB^{\text{EIF.}}$	0.880605	0.875953	0.878006	0.954259
4	**Ethereum**	$XGB^{\text{EIF.}}$	0.887876	0.888798	0.888324	0.958270
5	**Ethereum**	$XGB^{\text{EIF.}}$	0.888189	0.888189	0.888189	0.956110
Avg	**Ethereum**	$XGB^{\text{EIF.}}$	0.883306	0.882275	0.882730	0.953667

APPENDIX B: HISTOGRAMS OF IMPORTANT FEATURES

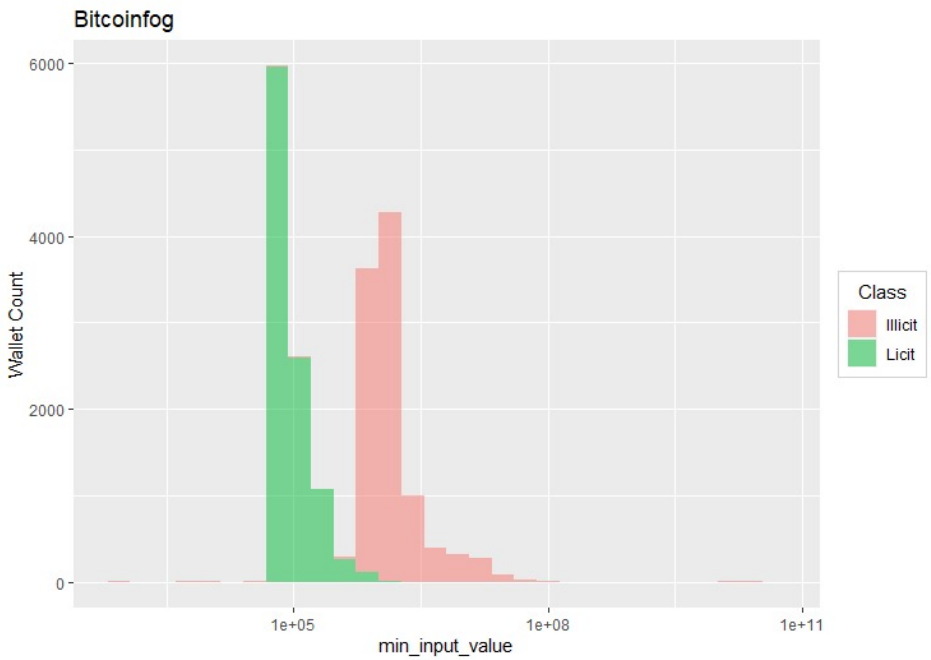

Figure B.1. Minimum Input Value Histogram - Bitcoinfog.

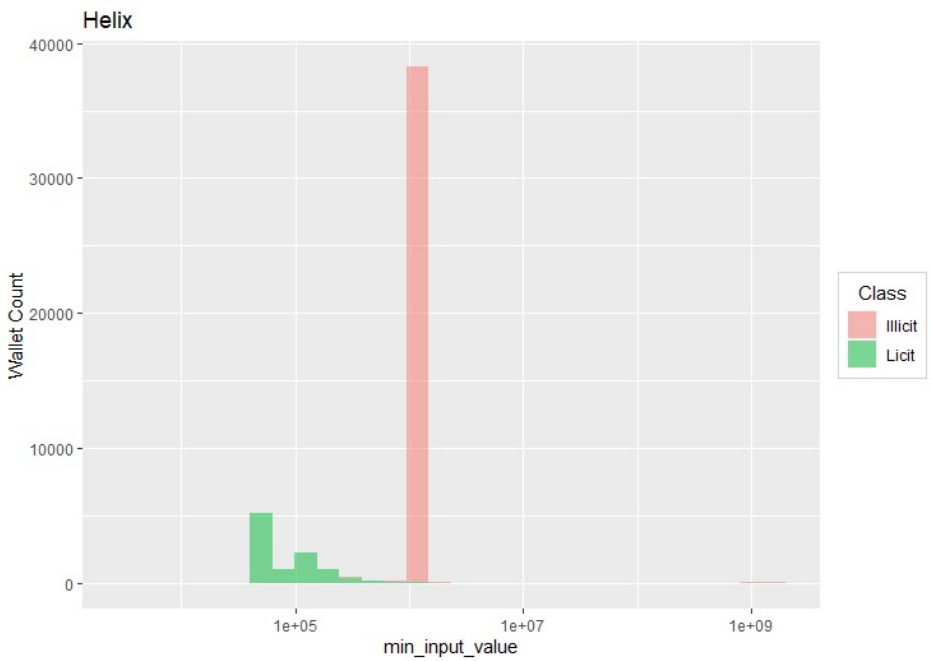

Figure B.2. Minimum Input Value Histogram - Helix.

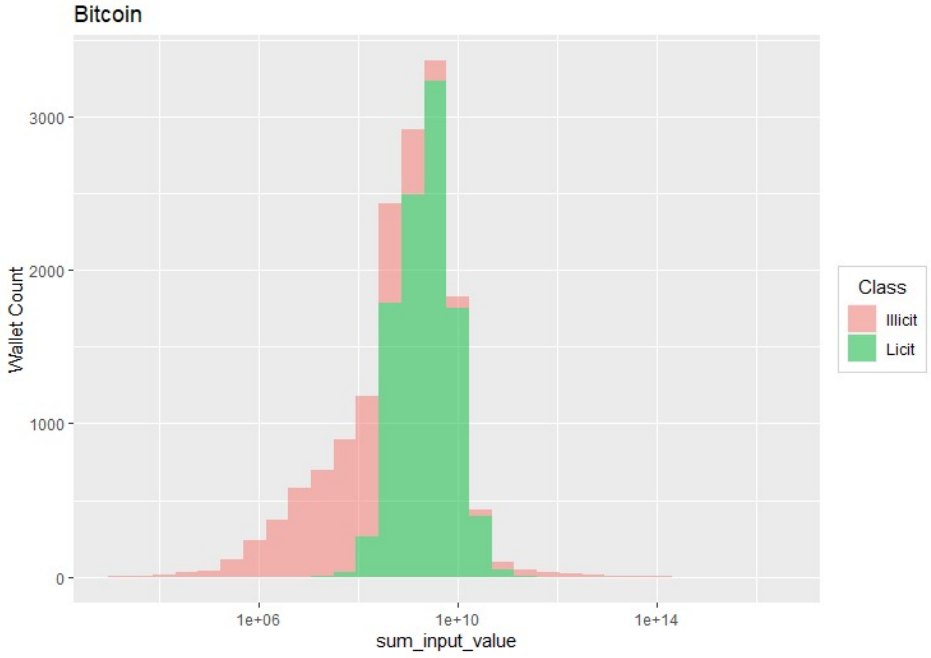

Figure B.3. Sum Input Value Histogram - BIllicit−Licit.

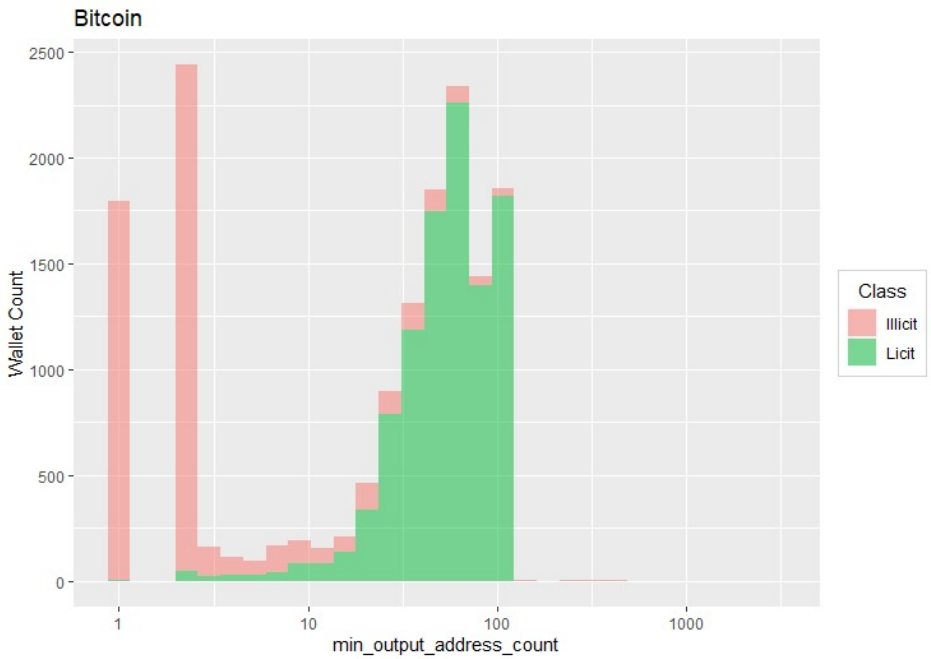

Figure B.4. Minimum Output Address Count Histogram - BIllicit−Licit.

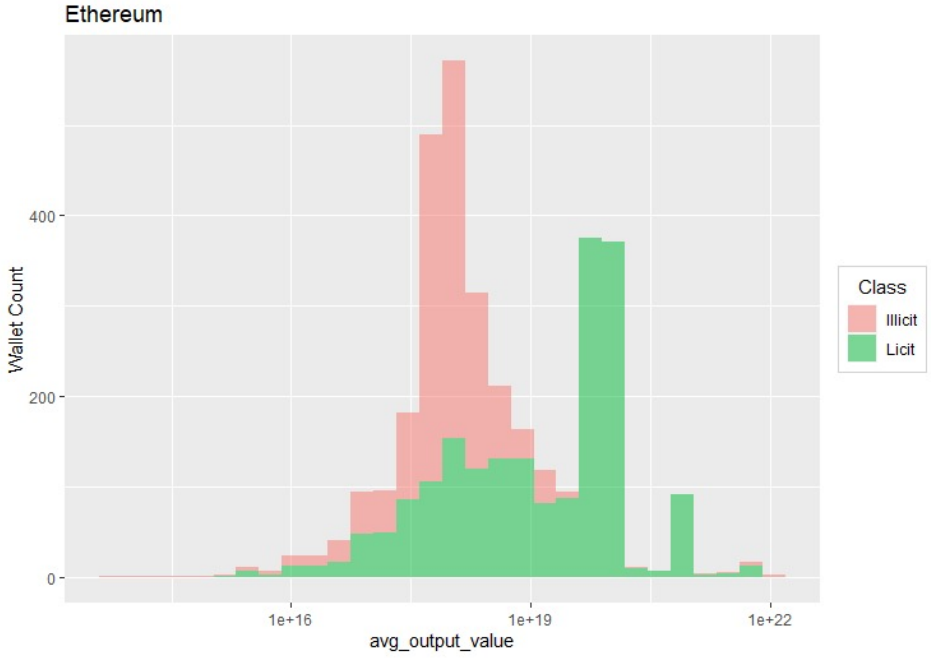

Figure B.5. Average Output Value - Ethereum.

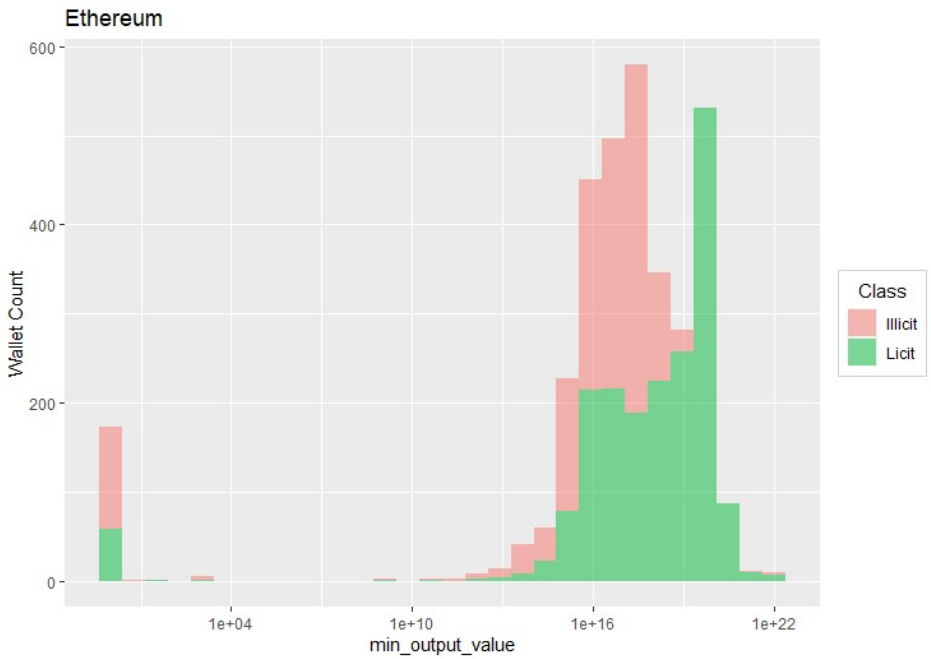

Figure B.6. Minimum Output Value - Ethereum.

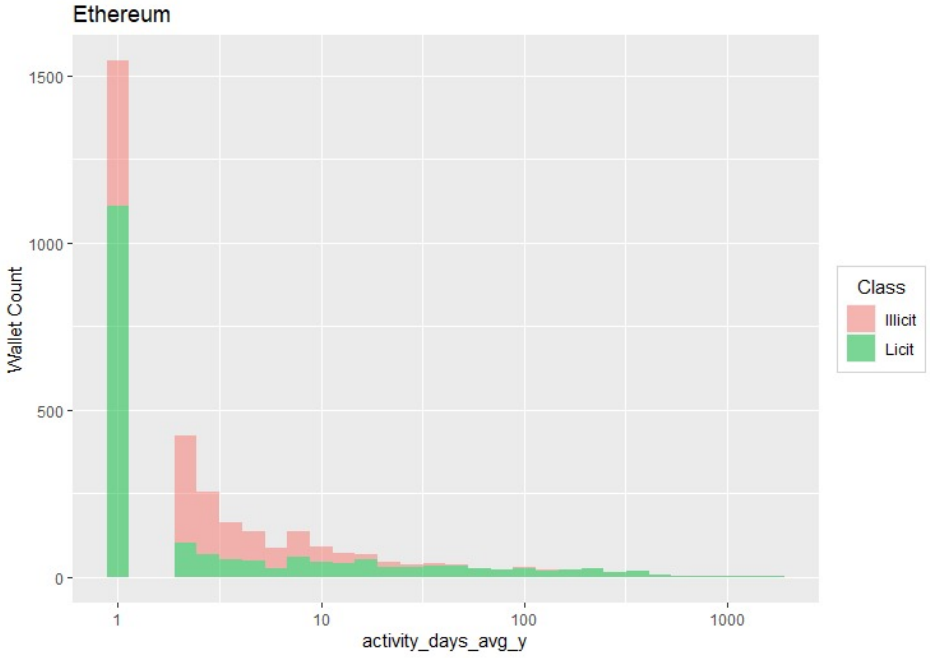

Figure B.7. Activity Days - Ethereum.

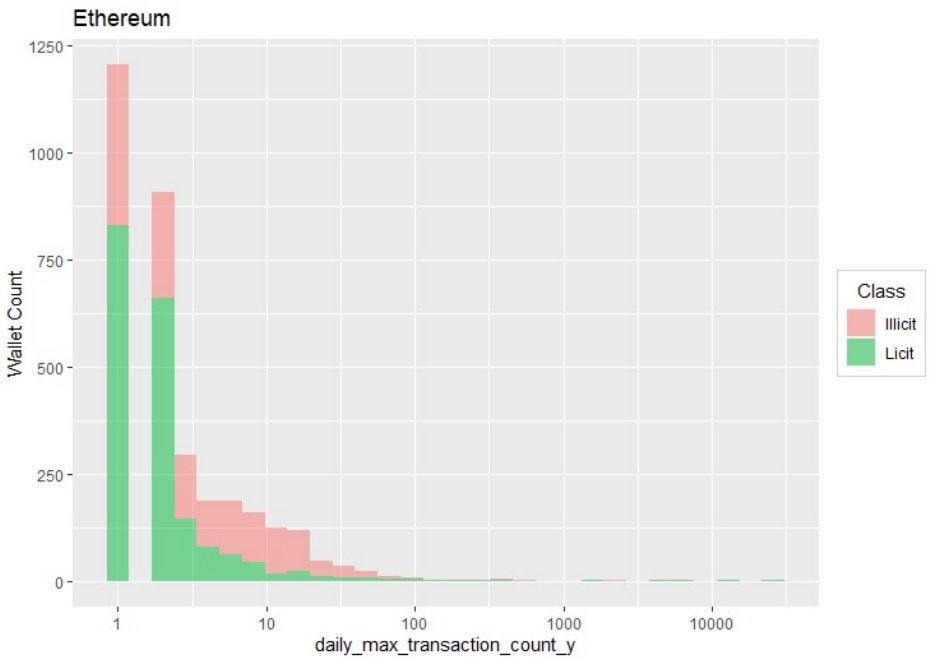

Figure B.8. Daily Maximum Transaction Count - Ethereum.