**Predicting Ethereum Fraudulency using**

**ChaosNet**

**Anurag Dutta 1[**0000 – 0002 – 5787 – 3860**] and Samrat Ray 2[**0000 – 0002 – 9845 – 2974**]**

1. Undergraduate, Department of Computer Science and Engineering, Government College of

Engineering and Textile Technology, Serampore, India

1. Research Scholar, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

**Abstract.** Cryptocurrencies are in high demand right now, perhaps because of their volatile nature and untraceable difficulties. Bitcoin, Ethereum, Dogecoin, and others are just a few. This research seeks to identify falsehoods and probable fraudulences in Ethereum transactional processes. We have provided this capability to ChaosNet, an Artificial Neural Network constructed using Generalized Luroth Series maps. At many spatiotemporal scales, Chaos has been objectively discovered in the brain. Several neuronal models, including the Hindmarsh-Rose neuron model, exhibit sophisticated Chaotic dynamics, and individual brain neurons are known to display chaotic bursting activity. Although Chaos is included in several Artificial Neural Networks (ANNs), such as Recurrent Neural Networks, no ANN exist for classical tasks that is fully made up of Chaoticity. ChaosNet uses the chaotic GLS neurons' topological transitivity property to perform classification problems with cutting-edge accuracy in the low training sample pool. This network can perform categorization tasks by learning from a limited amount of training data. ChaosNet utilizes some of the best traits of biological neural networks, which derive from the rich Chaotic activity of individual neurons, to solve difficult classification tasks on par with or better than standard Artificial Neural Networks. It has been shown to require much fewer training samples.

**Keywords:** Crypto – Currency • Blockchain • ChaosNet • GLS Neurons • Artificial Neural Network

# Introduction

Because of the advancement of Artificial Intelligence, learning via algorithms like Machine Learning (ML) and Deep Learning (DL) has gained appeal and is a hot topic, with applications in practically every sector of human endeavours. Among these are, to name a few, Voice Processing, Computer Vision, Cyber Security, and Medical Diagnosis. Despite being influenced by the biological brain, the learning and memory encoding processes in humans are not directly tied to these algorithms. These artificial neural networks' (ANNs) learning processes for changing weights and biases are based on optimization strategies and the minimizing of loss/error functions. As more and more new data is fed into the system, the ANNs currently use a huge number of hyperparameters that are fixed via an *ad hoc* approach for better prediction. These synaptic alterations are based primarily on facts and lack or have little solid theoretical support. Additionally, these methods need a huge quantity of training data to be able to accurately forecast or classify the distribution of the target classes.

ANNs have achieved great success, but when it comes to doing tasks like natural language processing, they fall well short of the human intellect. Researchers are concentrating on creating biologically inspired algorithms and architectures in order to utilize the remarkable learning capabilities of the human brain as well as to better understand the brain. This is being done in relation to memory encoding and learning. One of the brain's most intriguing traits is its capacity for "Chaos" – the phenomenon whereby straightforward deterministic nonlinear systems exhibit complex unexpected and random – like behavior. Electroencephalogram (EEG) signals are known to have chaotic dynamics. A neural system's sensitivity to little changes in internal functioning characteristics aids in producing the optimal response to various influences. This characteristic resembles the chaotic systems' dynamical characteristics. Furthermore, it is evident that the brain is constantly switching between several states rather than returning to homeostasis after a transient. For this reason, it is hypothesized that the brain can display a variety of behaviours, including periodic orbits, weak chaos, and strong chaos, depending on the functional parameters of the neurons. Brain networks, which are made up of billions of neurons, exhibit chaotic activity, but so do the dynamics of individual neurons at the cellular and subcellular levels. These neurons' ability to build impulse trains is what allows the brain to transmit and store information. When various ions pass across the axonal membrane and affect the voltage across it, impulses or action potentials are produced. For the interaction between the ion channels and the axon membrane, Hodgkin and Huxley were the first to put forth a dynamical system's model that is able to produce accurate action potentials. Later, it was suggested to use its streamlined counterparts, such as the Hindmarsh-Rose model and the Fitzugh-Nagumo model. These models all display chaotic behavior.

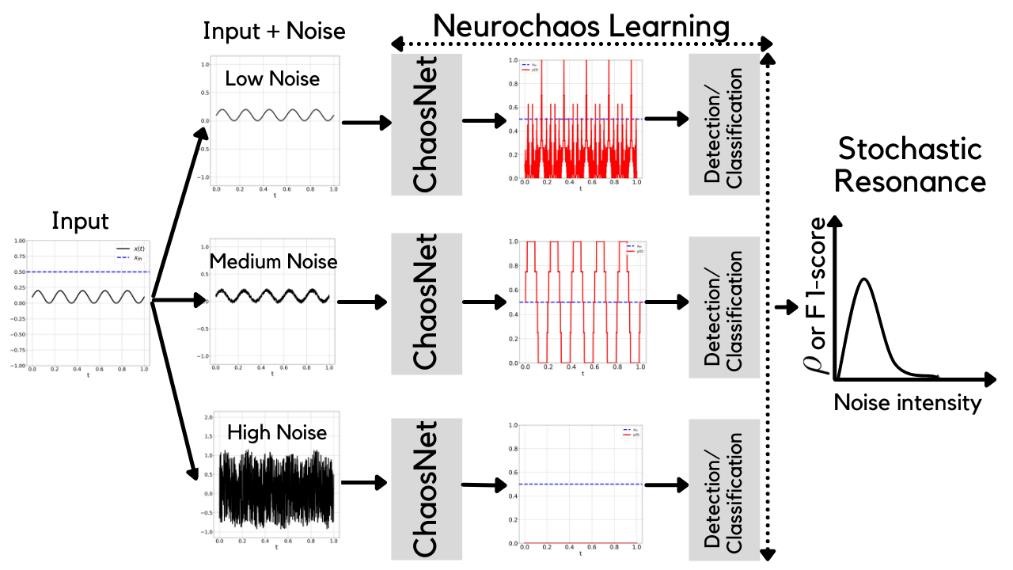
Recurrent neural networks are one type of artificial neural network that exhibits chaotic dynamics; however, as far as we are aware, none of the architectures proposed thus far for classification tasks show chaos at the level of individual neurons. However, other chaotic neuron models have been proposed as a theoretical explanation for memory encoding in the brain.

One of these models is the Aihara model, that has been applied to cognitive tasks in the network's erratic periodic orbits. Freeman, Kozma, and colleague developed chaotic simulations that were motivated by the mammalian sensory pathway to demonstrate the process of memorising scents. Chaos in neural networks has also been studied by Tsuda and others. Globally coupled chaotic maps' dynamical properties have been studied by Kaneko, who hypothesised that these networks would be able to handle biological data.

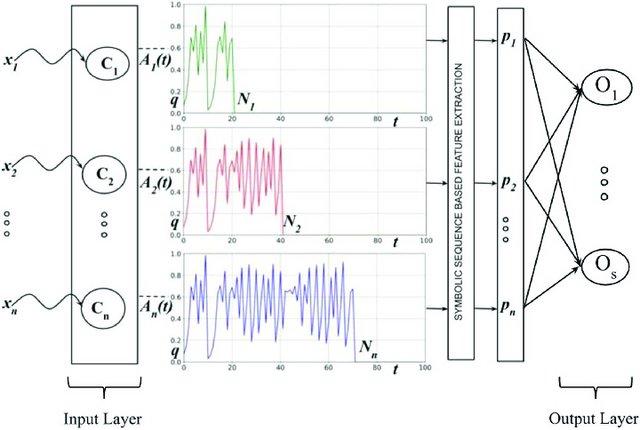
Generalized Luröth Series (GLS) 1D chaotic map neurons make up ChaosNet, an artificial neural network (ANN). This network can learn from a small number of training examples to perform classification tasks. To utilize some of the best characteristics of biological neural networks, ChaosNet was developed. It has been demonstrated that, while using significantly fewer training samples than traditional ANNs, it can perform difficult classification tasks on par with or better than conventional ANNs.

ChaosNet, which was inspired by biological neurons, uses a "spike-count rate"-like property of the firing of chaotic neurons as a neural code for learning. Additionally, the network can exhibit a hierarchical architecture that can incorporate information as it is transmitted to deeper, higher levels of the network. Generalized Luröth Series, or GLS, is a piecewise linear 1D chaotic map that represents the neuron that we suggest. Examples of GLS include the well-known Tent map, Binary map, and its skewed relatives. The sorts of GLS neurons that are employed in ChaosNet are

and



**Fig. 1.** Neurochaos Learning



**Fig. 2.** The architecture of ChaosNet Luroth neural networks for purposes relating to classification. 𝐶1, 𝐶2, . . . , 𝐶𝑛 are the unit dimensional GLS neurons. Each neuron initially exhibits 𝑞 units of normalised neuronal activity. The input to the network, or the normalized collection of stimuli, is denoted by the {𝑥𝑖}𝑛𝑖=1. When a GLS neuron 𝐶𝑖's chaotic activity value 𝐴𝑖(𝑡), starting from initial neural activity (𝑞), reaches the 𝜀 -neighbourhood of stimulus, it stops firing chaotically. This neuron has a “firing time” of 𝑁𝑖 ms. 𝐴𝑖(𝑡) contains topological transitivity symbolic sequence feature 𝑝𝑖. This feature is extracted from 𝐴𝑖(𝑡) of the 𝐶𝑖's GLS-neuron.

A cryptocurrency, often known as a crypto-currency or simply a "crypto," is a type of digital money that operates as a means of exchange over a computer network and is not supported or maintained by a single central organization, such a bank or government. It is a decentralized means of verifying that the parties to a transaction genuinely have the funds they claim to have, eliminating the need for traditional middlemen like banks when money is being transferred between two businesses. Digital ledgers, which are computerized databases that use safe encryption to protect transaction records, regulate the production of new currencies, and confirm ownership transfers, are used to keep individual coin ownership records. Cryptocurrency is typically not authorized by a centralized unit and isn't exist in tangible form like paper money. Cryptocurrencies frequently use decentralized control, in contrast to digital currencies governed by a central bank (CBDC). A cryptocurrency is typically seen as centralized when it is minted, generated in advance of issuance, or released by a single issuer. Each cryptocurrency operates using distributed ledger technology, often a blockchain, which acts as a public database of financial transactions when used with decentralized control. Currency, commodities, and stocks are traditional asset classes and macroeconomic indicators with moderate sensitivity to cryptocurrency returns.

Financial or personal gain is the intended outcome of cryptocurrency fraud, which is a dishonest behavior in the cryptocurrency business. By convincing their unwitting victims to take an action, such as clicking on a link or disclosing personal information, scammers, and hackers on the internet hope to make some quick cash.

5

For cryptocurrency scams, criminals frequently try to gain access to a victim's digital wallet in order to steal their cryptocurrency assets. Typically, they will ask you to connect your wallet to a bogus website or deceive you into giving them access to your wallet's private keys. Cryptocurrency Fraudulences can be of many types, like,

* *Phishing*: Even though fraudsters are nothing new, individuals continue to fall for them every day. A malicious hyperlink in an inbox or a bogus website that occasionally uncannily resembles its genuine counterpart can both be used in phishing scams. Your personal information, such as your internet passwords or the private keys to your crypto wallet, may be stolen using the link or website.

* *Middle Man attacks*: Man-in-the-middle assaults are a technique used by con artists to obtain your personal information, much to phishing scams. To access your bitcoin wallet or private account information, a guy will disrupt a Wi-Fi session on a broad network as opposed to doing so through links. Use a VPN to secure your data while depositing cryptocurrency to avoid this.

* *Investors’ Scam*: Investment managers that offer to help you make significant improvements on your portfolio may be posing as fraudsters. These dishonest people will entice customers to transmit them cryptocurrencies and may even promise to increase its worth by 50 times. Forbes Advisor does caution, though, that "if you comply with their demands, kiss goodbye to your cryptocurrency." With this scam, the con artist is probably deceiving several people, taking their cryptocurrency with them, and then vanishing.

* *Pump & Dump*: This is true for both regular stock markets and cryptocurrency marketplaces. When a coin launches, its owners sell all of their holdings, which is known as a pump and dump strategy. As a result, the price reaches an erroneous peak before dropping sharply after the initial public offering is over. False statements made about a project that cause a lot of hype can make these tactics worse.

# Ethereum

Ethereum is, at its core, a proof-of-stake decentralized global software platform. It is well known for its ether cryptocurrency (ETH). Anyone can use Ethereum to develop any safe digitizing. It has a currency designed to recompense users for work done in support of the blockchain, but if accepted, users may also use it to exchange for physical goods and services. Ethereum has the characteristics of being extensible, adaptable, anonymous, and decentralized. It is the decentralized cryptocurrency of choice for programmers and businesses building technology on top of it to alter multiple industries as well as how people go about their daily lives. Ethereum was first described in a white paper in late 2013 by Vitalik Buterin, a developer and cofounder at Bitcoin Magazine, as a mechanism to construct decentralized apps. Buterin claimed to the Bitcoin Kernel maintainers that Nature of blockchain technology may benefit from uses other than money and that a more sophisticated language for developing applications was required. In early 2014, a Swiss company, Ethereum Switzerland GmbH, began formal development of the software underlying Ethereum (EthSuisse). Before it could be implemented in software, the idea of storing executable smart contracts in the blockchain had to be specified. This work was done in the Ethereum Virtual Machine specification by Gavin Wood, the Ethereum Yellow Paper's then-chief technical officer. Following that, the Ethereum Foundation (Stiftung Ethereum) was established as a Swiss non-profit organization. From July through August 2014, an online public crowd sale was held in which individuals purchased the Ethereum value token (ether) with another digital currency, bitcoin. Although the technical innovations of Ethereum were first praised, questions about its scalability and security were raised. In order to build and attain consensus on an ever-expanding collection of "blocks," or groupings of transactions, known as a blockchain, Ethereum is an epicondyle, or virtual collective, of computer nodes. For the sequence that must come before each block in order for it to be considered authentic, each block has a distinct identifier. When a base station adds a block to its chain, it executes the actions in the block in the designated order, each of which has the potential to alter the ETH balance and other rack values of Ethereum accounts. In a Merkle tree, the "state," or collection of these totals and values, is held on the node apart from the blockchain. Only a limited portion of the network, known as its "peers," are accessible to each node. Every time a node wants to add a new transaction to the chain, it sends copies of the transaction to all of its contemporaries, who then send copies to all of their contemporaries, and so forth. It spreads throughout the network in this way. All these fresh transactions are kept track of by a group of nodes known as miners, who use them to build new blocks and distribute them to the remainder of the network. Every time a node receives a block, it verifies the validity of the block and of each transaction contained inside. If the block is found to be valid, it is added to the blockchain and each transaction is carried out. A node may receive numerous blocks that are vying to succeed a specific block since block generation and broadcasting are permissionless. The node records each valid chain that results from this and routinely discards the shortest one: The Ethereum protocol states that the longest chain is to be taken into consideration at any given time.

# Dataset Description

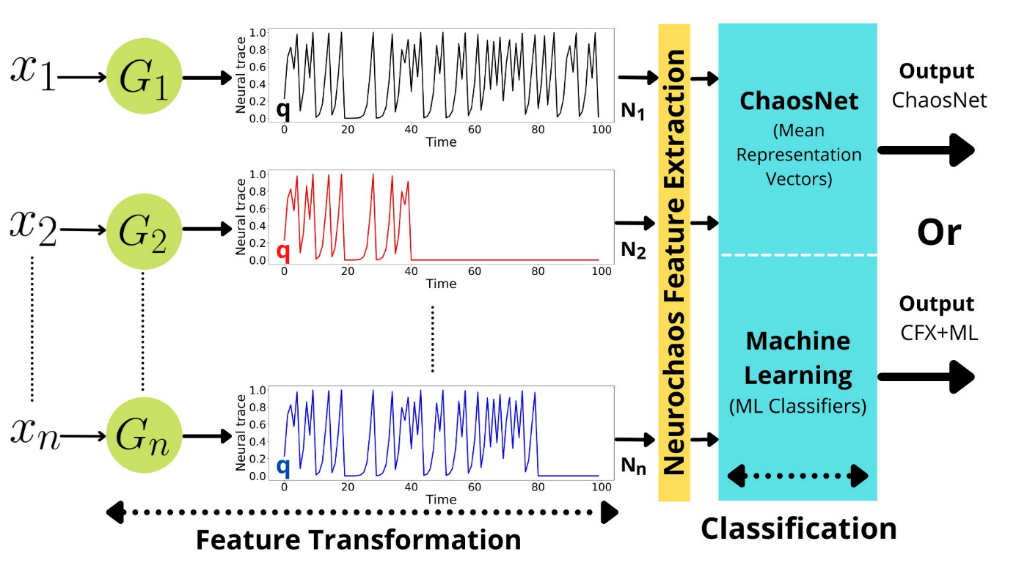
We have collected a set of Ethereum Transaction Details using Etherscan API, and ehterscamdb API. The Dataset, has 14 Features namely,

* Avgminbetweensenttnx: Minutes between each transaction on average for the account.
* Avgminbetweenreceivedtnx: Minutes between transactions received on average for the account.
* TimeDiffbetweenfirstand\_last(Mins): Between the first and last transaction, in terms of time.
* Sent\_tnx: Total volume of typical transactions sent.
* Received\_tnx: Total volume of typical transactions received.
* NumberofCreated\_Contracts: Total number of contract transactions created
* UniqueReceivedFrom\_Addresses: Total unique addresses from which transactions were sent to the account
* UniqueSentTo\_Addresses: Total unique addresses from which transactions were sent from the account
* MinValueReceived: Lowest amount of Ether ever received
* MaxValueReceived: Highest amount ever paid in Ether
* AvgValueReceived: Ever received an average amount in Ether
* MinValSent: The smallest amount of ether ever sent
* MaxValSent: Highest amount of ether ever transferred
* AvgValSent: Average amount of ether transmitted over time

The dataset have been maid available online at <https://github.com/Anurag-Dutta/Ethereum_Trasnsaction/blob/cd6508c3c13ca7292b02e333d8e3272965d3d88c/Eth_Pub_19122022.csv>.

# ChaosFeatureEXtractor + ML Classifiers

Using ChaosNet Standalone, the performance is okay, but we can do better if we make use of any better ML Classifier as a conjunction to the Chaos Feature Extractor.



**Fig. 3.** Architecture proposing Conjunction of the Chaos Feature Extractor with standard ML Classifiers. The three actions involved includes Feature transformation, feature extraction from Neurochaos, and classification are the first two steps. ChaosNet or any other ML classifier could be selected as the chosen classifier.

ChaosNet uses 3 hyperparameters

1. INA - Initial Neural Activity
2. EPSILON\_1 - Noise Intensity
3. DT - Discrimination Threshold

The memory of this Single Internal Neuron is corresponding to the Initial Neural Activity. As individual Machine Learning Classifiers, we have used AdaBoost, and kNN (k Nearest Neighbours). Also, we have made use of ChaosNet Standalone.

The respective values of the hyperparameters for the same was tuned to

INITIAL\_NEURAL\_ACTIVITY = [0.38]

DISCRIMINATION\_THRESHOLD = [0.06]

EPSILON = [0.29]

for Standalone ChaosNet.

INITIAL\_NEURAL\_ACTIVITY = [0.36]

DISCRIMINATION\_THRESHOLD = [0.06]

EPSILON = [0.29]

for ChaosNet Feature Extractor conjugated with AdaBoost.

Yoav Freund and Robert Schapire developed the statistical classification meta-algorithm known as AdaBoost in 1995. They were awarded the 2003 Gödel Prize for their work. Its performance can be enhanced by combining it with a variety of other learning methods. The results of the other learning algorithms, or "weak learners," are merged to create a weighted total that represents the boosted classifier's final results. Although AdaBoost can be used to many classes or bounded intervals on the real line, it is often shown for binary classification. AdaBoost is adaptive in the sense that it modifies succeeding weak learners in favour of examples that were incorrectly identified by earlier classifiers. It may be less prone to the overfitting issue than other learning algorithms in particular situations. It can be demonstrated that the final model converges to a strong learner even if the performance of each individual learner is just marginally better than random guessing.

INITIAL\_NEURAL\_ACTIVITY = [0.039]

DISCRIMINATION\_THRESHOLD = [0.070]

EPSILON = [0.023]

for ChaosNet Feature Extractor conjugated with k Nearest Neighbours.

One of the simplest supervised learning-based nonparametric machine learning algorithms is K-Nearest Neighbour. Assuming that new cases and data are similar to existing cases, classifying new cases into categories that are most similar to existing categories, storing all relevant data, and based on similarity putting fresh data into categories. Therefore, it is simple to categorise new data into the relevant categories using the K-NN method. Although K-NN algorithms can be applied to classification and regression problems, they are most frequently utilised for classification issues. In other words, no presumptions regarding the underlying data are made. It is also known as a delayed learning algorithm since it saves the dataset and modifies it during classification rather than instantly learning from the training set. The k-NN algorithm only keeps the training phase dataset, and it classifies fresh data into categories that are like the new data as it comes in.

In order to evaluate performance, we used the macro F1-score. The F1 score can be thought of as a harmonic mean of precision and recall, with the best value being 1 and the poorest being 0. Precision and recall both contribute equally to the F1 score; "macro" calculates the metrics for each label and determines their unweighted mean. Label imbalance is not considered in this. The confusion matrix is used to calculate this measure. Mathematically,

where,

|  |  |
| --- | --- |
| *Algorithm* | *Macro F1 Score (Training)* |
| ChaosNet Standalone | 0.5802753655203908 |
| Chaos Feature Extractor + AdaBoost | 0.8125910159305623 |
| Chaos Feature Extractor + kNN | 0.7937217353400664 |

|  |  |
| --- | --- |
| *Algorithm* | *Macro F1 Score (Testing)* |
| ChaosNet Standalone | 0.5752543039000217 |
| Chaos Feature Extractor + AdaBoost | 0.6649360740269832 |
| Chaos Feature Extractor + kNN | 0.7888128840520701 |

# Conclusion

In the field of machine learning, making decisions when unusual events are present is a difficult task. This is because unusual occurrences have few data examples, and the issue is ultimately one of imbalanced learning. In this work, we have taken use of Neurochaos Learning (NL) architectures' usage of ChaosFEX (CFX) feature modification for imbalanced learning. Since, Cryptocurrencies are in it’s very nascent stage currently, and due to their masking nature, finding data involving transaction in the Ethereum is quite difficult, and not much can be obtained from that. Like, even if we manage to snoop into the transaction details, the details won’t be necessarily sufficient to fulfil the requirements of the classical ML Classifier Algorithms. In this work, we have tried to make use of ChaosNet along with it’s indigenous Feature Extractor to try out the prediction of possible fraud in the Ethereum Transaction. We made use of the Standalone ChaosNet, that gave us a F1 Score of 0.58 for Training, and 0.57 for Testing, which isn’t a good hold. Further, we have made use of the ChaosNet Feature Extractor assisted with Adaptive Boosting to get a F1 Score of 0.81 for Training, and 0.66 for Testing. Finally, we made use of the ChaosNet Feature Extractor assisted with *k* – Nearest Neighbours to get a F1 Score of 0.79 for Training, and 0.78 for Testing, which is the maximum we can get of it. So, we can conclude of by saying that the ChaosNet Feature Extractor assisted with *k* – Nearest Neighbours is the best for Predicting Possible Fraudulences in the Ethereum Transaction Dataset. Notably, the F1 Scores used here, are the Macro F1 Score.

Data, code, and supporting materials are publicly available via https://github.com/AnuragDutta/BrainTumor publicly available.