Cryptocurrency Price Prediction Using Machine Learning

K. Ramya Laxmi*[1] , Marri Abhinandhan Reddy[2] , CH. Shivasai[3] , P. Sandeep Reddy[4]

[1]   AssistantProfessor, Department of CSE, Sreyas Institute of Engineering and Technology, Nagole, Hyderabad, India
      e-mail : kunta.ramya@gmail.com
[2-4] Student, Department of CSE, Sreyas Institute of Engineering and Technology, Nagole, Hyderabad, India.

Abstract

The digital currency in which encryption techniques are used to regulate the generation of units of currency is said to be called cryptocurrency. The technology used here is used to explores the next day change in the price of cryptocurrency. It is a challenge for a common person to achieve with varying degrees of success. But this is achieved through the implementation of a optimized recurrent neural network (RNN) and a Long Short Term Memory (LSTM) network.

Keywords: Recurrent Neural Network, Long Short Term Memory, Machine Learning, Cryptocurrency.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, (2020); DOI : 10.18090/samriddhi.v12iS3.3

Introduction

Crypto currencies, to a great extent "Bit coin", have been one of the top hit in social media and in technical world. Their high volatility leads to the great potential of high profit if intelligent inventing strategies are taken. It seems that every one in the world suddenly start to talk about Crypto currencies. Unfortunately, due to their lack of indexes, Crypto currencies are relatively unpredictable compared to traditional financial instruments. Crypto currency is a way to help the market become healthier and more transparent. This technology is indeed having challenges to overcome. The later section gives more knowledge on how to predict the price of these Crypto currencies with Deep Learning. We choose bit coin as an example from different number of crypto currencies [1-5].

Literature Survey

In step with privacy, safety and agree with the concepts of Cloud Computing, the author states that, Cloud computing refers to a basic infrastructure for an upcoming service model that offers low cost service with the help of shared computing and garage resources, mixed with an on-call for provisioning mechanism counting on a pay-regular with-use organisation version. Impact of such functions is higher in case of IT (information technology) budgeting

however additionally have an effect on traditional safety, don't forget and privateness procedure. Various benefits of cloud computing are its rapid scaling capability, save facts remotely and dynamic sharing of offerings come to be dangers in maintaining a degree of warranty sufficient to sustain self belief in potential customers. Some core conventional mechanisms for addressing privateness (including version contracts) aren't bendy or dynamically sufficient, hence new approaches are required to evolve the new paradigm in shape. During such financial disaster, the troubles related to the trust, protection and privateness that arises within the cloud computing framework and their communication models are addressed.

Cloud Computing trust Procedure is considered to be a crucial trouble in cloud computing in gift workout it depends in massive element on reputational belief and self-assessment through cloud service carriers. This paper begins with a survey of contemporary procedures for organizing take delivery of as real with, and comment on their barriers. We then cope with those obstacles through manner of providing extra rigorous mechanisms primarily based totally on proof, certification of attribute certification, and validation, and the paper concludes via suggesting a framework for integrating diverse remember procedures together to show chains of trust inside the cloud [1-9].

PROPOSED METHODOLOGY AND SOLUTIONS

Crypto currency price prediction using machine learning and this project has got some real time analysis and prediction power by integrating Machine learning as its core thing. The process of prediction of prices using machine learning as follows

Data

First data is pushed up over the stack because of its priority as an object. It is the main object for training our machine learning model. There's a lot of data related to Bitcoin about 37 characteristics of it n "bitcoin.com" . This data has been collected since July of 2010, hence it makes availablearound 60 thousand data points to be processed. With this large amount of data, there exists a way to look into the possibility and probability of predicting future prices.

Machine Learning

Recently, Machine learninghas emerged into a promising technique due to increase in the availability of data as well as computational power. The machine learning has a set of neurons connected like a neural networks which serves as an artificial brain, hence it is possible for the machines to find patterns for huge datasets with minimum human annotation. Researches are being carried out to construct complex network with multiple layers which can handle much difficult problems. Machine learning has a lot of applications in various fields. Recent applications of machine learning includesfacial recognition,self-driving cars and language translation.

Such a power was the best bet to see if It could predict cryptocurrency prices. A simple Architecture of neural network [10, 11].
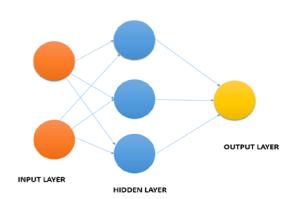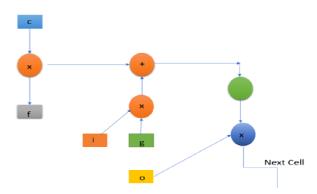
Figure 1 : Simple feedforward neural network

LSTM Cell

LSTM is a kind of neural network that paves a revolutionary way in finding patterns. RNN is compoed of a large number of LSTM cells. The question arises in the working of the RNN. It utilises a set of specialised gates for allowing each layer of LSTM to get information from both previous as well as the current layers. The data flows among number of gates as well as different activation functions and it flows through LSTM cell. In this, each LSTM cell is allowed to remember the pattern related to necessary data for a short span of time and forget unnecessary data.

Functions (i.e. the tanh function) and is passed throughout the LSTM cells. The main advantage of this is that it allows each LSTM cell to remember patterns for a certain amount of time — they essentially can "remember" important information and "forget" irrelevant information.

Figure 2. LSTM cell and its internal components — cell state c. input gate i, output gate o, forget gate gand external input gate g

Table-1 : Details of Data Collection

	Date	Open	High	Low	Close	Volume	Market Cap
0	2017-11-19	7766.03	8101.91	7694.10	8036.49	3149320000	129595000000
1	2017-11-18	7697.21	7884.99	7463.44	7790.15	3667190000	128425000000
2	2017-11-17	7853.57	8004.59	7561.09	7708.99	4651670000	131026000000
3	2017-11-16	7323.24	7967.38	7176.58	7871.69	5123810000	122164000000

The different phase that need to be undergo during the process of predicting the prices of crypto currencies are
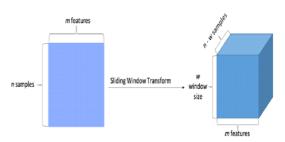
- Data Collection
- Data processing
- Data training

Data collection

In Data collection the massive data is collected and stored for performing data filtering operations . The first step for starting the process. The data collection phase begins with downloading all of the data available online, the downloaded data consists 37 different Excel files. There was a way to feed each file separately into the neural network, but it was way easier to manually combine the files into one big file. So by collecting a huge file that has more columns by around more rows. Here you can see how data is collected in csv files which includes different parameters Market Price, Open, Close, High, Low, Volume, Market Cap.

Data Preprocessing

Data preprocessing has some pretty technical steps to it. The first thing to do is the application of transformation of sliding window to data. It begins with sliding a imaginary window over the big Excel file for making it to arrays of days by feature. So, 3D rectangular prism is formed by transforming a 2D rectangle. The next step performs some normalization on the data following splitting the data for training, validation and testing. Variations are observed in the range of values and it utilises 10 percent data for testing and the rest of the data is available for training.

Figure 3: Representation of how the sliding window processing changes the data

Here the data is the above data collected from (.CSV) is loaded as data set matrices and normalization is performed and there after it goes with splitting of data and improve them by training. After a while of training it should be reshaped as X and Y wher x = t and Y = t+1.As there are different types of models with different features we need to select a Models which does major amount of work and suits our way of solving a problem so we use LSTM model.

RESULTS AND DISCUSSION

Deep Learning Model

To make the model learn we need to focus on LSTM & RNN to allow identification of smaller sequence patterned data and the price of the next day is predicted. A python library known as"Keras" utilized for the creation of neural network uses TensorFlow backend.

The layers are as follows:

- Input layer (takes data of shape n samples x 50 x 37)
- Bidirectional LSTM layer (returns a sequence, 100 cells)
- Dropout layer (20% dropout — reduces overfitting)
- Bidirectional LSTM layer (returns a sequence, 100 cells)
- Dropout layer (20% dropout — reduces overfitting)

- Bidirectional LSTM layer (doesn't return a sequence, 50 cells)
- Output layer (returns the predicted next day price of Bitcoin)

Training

The training of the data needs a large amount of filtered data sets. When data reaches a stabilization point it doesn't increases it results efficiency so by using some function we can improve it by loss function, activation function, optimizer.
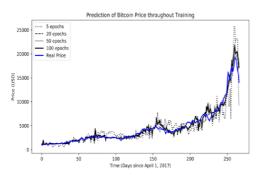

Figure 4: Progress of the model throughout training (more epochs = more iterations)

Linear Regression

The main concept in our project is prediction and so to build a relationship between those models we use Linear Regression to draw a relationships between those data objects. Linear relationship among target as well as predictors is found usingLinear regression.

When it is possible to express a variable accurately by the other, then the relationship between them are known as deterministic. For instance, temperature can be predicted in terms of Fahrenheit when it is provided in degree Celsius. Relationship among variables cannot be determined accurately using statistical relationship. The relationship between weight as well as height can be considered as an example.

The main objective is to attain a line that best fits the data. A line is said to be best fit when it possesses a smaller total prediction error. Distance between points in regression line is called as error.

CONCLUSION AND FUTURE SCOPE

The results always relays only on the training data sets we can cross check the results with present prices by statistical significance test. We totally

can't depend on this model because some time it prone to false prediction.

REFERENCE

[1] Sheikh MahbubHabib, Sebastian Ries, Max Muhlhauser, Towards a Trust Management System for Cloud Computing, IEEE Trust,Security and Privacy in computing and Communications (TrustCom), Pages 933-939, 2013

[2] B.Kezia Rani, Dr.B.Padmaja Rani, Dr. A. VinayaBabu, Cloud Computing and Inter-Clouds Types, Topologies and Research Issues,ELSEVIER, Volume 50,Pages 24-29, 2015

[3] Sheikh MahbubHabib, SaschaHauke, Sebastian Ries and Max Muhlhauser, Trust as a Facilitator in Cloud Computing: A survey,Journal of Cloud Computing,1:19,2012

[4] RajkumarBuyya, Christian Vecchiola, ThamaraiSelvi, Mastering in Cloud Computing, Morgan Kaufmann, May 2013.

[5] Maricela-Georgiana Avram, Advantages and challenges of adopting cloud computing from an enterprise perspective, ELSEVIER,Volume 12, Pages 529-534, 2014

[6] Soon-Keow Chong, JemalAbawajy, Masitah Ahmad, IsredzaRahmi, Enhancing Trust Management in Cloud Environment,ELSEVIER, Volume 129, Pages 314-321, 2014

[7] Dawei Sun, Guran Chang, Lina Sun, Xingwei Wang, Surveying and Analyzing Security, Privacy and Trust Issues in Cloud ComputingEnvironments, ELSEVIER, Volume 15, Pages 2852-2856, 2011

[8] Khald M. Khan and QutalbahMalluhi, Establishing Trust in cloud computing, Qatar University, IEEE IT professional, Volume 12(5), 2010

[9] Talal H. Noor and Quan Z. Sheng, Trust as Service: A framework for Trust Management in Cloud Environments, Springer, Volume6997, Pages 314-321,2011.

[10] Rohit Raja, Tilendra Shishir Sinha, Raj Kumar Patra and Shrikant Tiwari(2018), Physiological Trait Based Biometrical Authentication of Human-Face Using LGXP and ANN Techniques, Int. J. of Information and Computer Security, Vol. 10, Nos. 2/3, pp. 303- 320.

[11] Rohit Raja, Tilendra Shishir Sinha, Ravi Prakash Dubey (2016), Orientation Calculation of human Face Using Symbolic techniques and ANFIS, Published in International Journal of Engineering and Future Technology, Vol. 7, Iss.7, pp. 37-50, ISSN: 2455-6432.

[12] Rehmat Khan, Rohit Raja (2016) Introducing L1-Sparse Representation Classification for facial expression, Published in Imperial Journal of Interdisciplinary Research (IJIR), Vol. 2, Iss. 4, pp. 115-122, ISSN: 2454-1362.

[13] Rohit Raja, Tilendra Shishir Sinha, Ravi Prakash Dubey (2015), Recognition of human-face from side-view using progressive switching pattern and soft-computing technique, Association for the Advancement of Modelling and Simulation Techniques in Enterprises, Advance B, Vol. 58, N 1, pp. 14-34, ISSN:-1240-4543.