Adv. Math. Fin. App., 2022, **7(**1), P. 37-48

Advances in Mathematical Finance & Applications
www.amfa.iau-arak.ac.ir
Print ISSN: 2538-5569
Online ISSN: 2645-4610
Doi: 10.22034/AMFA.2021.1932547.1603

Research Paper

Capsule Network Regression Using Information Measures: An Application in Bitcoin Market

Mahsa Tavakoli[a], Hassan Doosti[b], Christophe Chesneau[c]

[a]*Independent Researcher, Mashhad, Iran*
[b]*Macquarie University, Sydney, Australia*
[c]*Universite de Caen-LMNO, Caen, France*

ARTICLE INFO
Article history:
Received 2021-06-06
Accepted 2021-09-05

Keywords:
Deep Learning
Financial Market
Prediction

ABSTRACT

Predicting financial markets has always been one of the most challenging issues, attracting the attention of many investors and researchers. In this regard, deep learning methods have been used a lot recently. Due to the desired results, such networks are always in development and progress. One of the networks that is being implemented in various fields is capsule network.This relatively new network uses structures called capsules, which makes better results than competing networks such as Convolutional networks in the problems such as image processing. The first time the classification capsule network was introduced, it was able to attract a lot of attention with its success on MNIST [1]data. In such networks, as in the other ones, the parameters are obtained by minimizing a loss function. In this paper, we first change the classification capsule network to a regression capsule network by modifying the last layer of the network. Then we use different information measures such as Kullnack-Leibler, Lin-Wang and Triangular information measures as a loss function, and compare their results with wellknown models including Artificial Neural Network (ANN), Convolutional Network (CNN) and Long Short-Term Memory (LSTM) as well as common used loss functions such as Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE). Using appropriate accuracy metrics, it is shown that the capsule network using triangular information measure is well able to predict the price of bitcoin for the medium and long term period including 10, 90 and 180 days.So its forecast accuracy is 94% in the long term and 64% in the medium term.

1 Introduction

Forecasting financial markets is always a challenge due to the noise and the amount of information available[1],[2]. With the existence of economic, political, and psychological variables, financial mar-

* Corresponding author. Tel.: +98 9155218339
E-mail address: *tavakolimtn@gmail.com*

kets usually have a non-linear and chaotic trend[3],[4]. However, recent studies have shown that patterns for prediction can be derived using appropriate forecasting tools, which in itself can lead to opportunities for profitability and risk management. There are a variety of methods to predict financial markets. Machine learning techniques are one of the most effective of these methods[5]. Indeed, pattern identification based on historical data using machine learning tools has been very successful[6],[7],[4],[8],[9]. Although other forecasting methods, such as statistical methods, also have acceptable performance in predicting financial markets, many studies, which compared machine learning methods with statistical methods, indicated that machine learning methods are more powerful than statistical models for high-frequency data[10].It is a common belief that machine learning algorithms are able to identify patterns in a data set [11].Recently, significant advances have been made in machine learning methods and algorithms, so that a field such as deep learning on the subject of prediction has received attention[12],[13],[14],[15],[16],[17],[18]. Deep learning is an advanced version of machine learning that has the ability to extract more complex information and discover deep nonlinear relationships and has therefore recently established itself in various fields, such as finance [5]. Various popular machine learning algorithms, including Artificial Neural Network (ANN), Support Vector Machines (SVM),Recurrent Neural Network(RNN), Long Short-Term Memory(LSTM), Convolutional NN (CNN), and random forest models have been widely used in price prediction[19],[20],[10].Depending on the application and the problem at hand, each of these methods has had some success[21],[22].On the other hand, since the first introduction of capsule networks and their success with MNIST data, it seems to have the potential to be successful for a variety of issues. As [23]has demonstrated, using the Capsule net improves the performance of stock movement prediction. In this paper, we use regression capsule networks that use various types of the loss function to forecast the price of bitcoin in the medium and long term and compare the results with other networks. The section of this article is as follows:
In the second section, the financial market forecasting literature using machine learning, especially deep learning, is presented. In the third section, we provide an overview of capsule networks. In the forth section, we define some information measures to use them as loss functions in our findings. The fifth section, which is the forecasting and modeling section, includes 5 sub-sections in which we provide details of modeling from data and methodology to compare the results of the model. The final section is devoted to summarizing and concluding the analysis.

2 Related works

There are much and varied research related to forecasting financial markets using machine learning methods. In this section, we summarize some of this research. In the first part of this section, we discuss some of the work done on forecasting with deep learning methods, and in the second section, we explain about some research regarding bitcoin market forecasting using deep learning.
Goodfellow et al. [24]in one study concluded that deep learning is a powerful tool in the context of the financial market. Fischer and Krauss[25]found that, in theory, deep learning tools, especially LSTM, are very suitable for analyzing multidimensional time series due to their long-term dependencies. Korczak and Hemes[26]showed that in the FOREX exchange market, CNN compared to MLP has a significant rate of return. Hu et al.[27] built a hybrid model using LSTM, and by importing news reports, they were able to increase the accuracy of exchange market forecasting. Fischer and Krauss [28] applied LSTM to forecasting the movements of prices in the S&P 500 stocks. They concluded that LSTM networks outperform memory-free networks in terms of accuracy. Gonçalves et al. [29]. Using various

deep learning techniques in the foreign exchange market showed that although LSTM seems theoretically more predictable, CNN works better than it in practice.[30]introduced a new function that showed it performed better than its well-known competitors. The purpose of introducing this function was to gain more profitability in the Tehran Stock Exchange trading. They introduced this function as an efficient function in practice. Research on financial market prediction using deep learning is very diverse and numerous. For example, one can refer to papers such as[31-35]. Much work has been done on bitcoin forecasting using deep learning, some of which we will mention below.

[36]used the ANN and SVM methods to predict the bitcoin market and concluded that using SVM traders could obtain conservative returns by adjusting risk.[37]studied various types of learning methods, including deep neural networks, convolutional neural networks, long short-term memory, and so on for prediction of bitcoin price and concluded that among regression-based models, LSTM works better, but in classification, DNN outperformed the other prediction models. McNally et al. [12]compared machine learning methods, including RNN and LSTM, with statistical methods such as autoregressive integrated moving average (ARIMA). They indicated that LSTM performs better than ARIMA. Jang and Lee [14], in their studies, showed the efficiency of Model Bayesian Neural Network (BNN) compared to support vector regression and linear regression. Jang et al.[38] compared the LSTM model with linear regression, SVR and ANN, and showed that a rolling window LSTM model outperforms the others. Also, Shintate and Pichl[16] introduced a new model called random sampling and indicated that it worked better than LSTM. Kim et al. [39] and Li et al.[40] predicted bitcoin fluctuations using social data. There are papers that investigate the introduced machine learning techniques, such as k-means or vector quantization with various loss functions[41],[42],[43],[44]. In this paper, we intend to use different information measures as a loss function in deep learning algorithms and use them to predict.

3 Capsule Network

Capsule network was introduced to cover some of the weaknesses of CNN and was also able to show successful performance on MNIST data. We first give an overview of the capsule network structure in the main paper [42] and then explain the changes we have made to it for the purpose of the bitcoin market forecast.

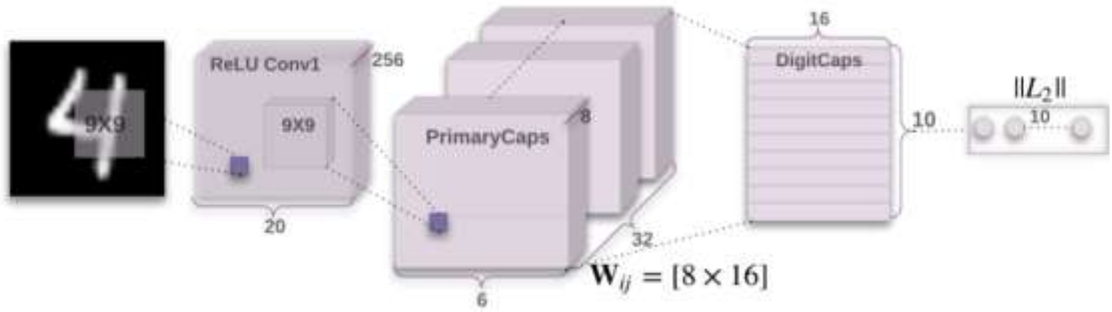

Fig. 1: Structure of Capsule Network Extracted From[42]

It can be seen in Figure 1, taken from the main paper. The architecture of the main capsule network is as follows: The first layer is ordinary convolutional with 256, 9×9 kernels. In the second layer, called

Primary Caps, each capsule contains eight convolution units with a 9×9 kernel. At this point, the capsules are multiplied by a weight matrix[2]to form the digit cap layer. Squashing and dynamic routing operations are performed in this layer. After the routing algorithm, the length of the output capsules is used for classification results. One can see [45] for more details. The following figure shows the general structure of the capsule Network that was implemented on MNIST data. In our problem where the output is a continuous vector instead of a class, and we are actually dealing with a regression problem, we proceed with the capsule architecture as follows: After completing the dynamic routing process, we use a dense layers an output layer with a neuron. This layer can extract the regression result.

4 Information Measures as Loss Function

In all optimization algorithms, the parameters are determined based on the loss function so that the loss function is minimized. Machine learning algorithms are no exception. In classification problems, the well-known loss function cross-entropy, and in regression problems, functions such as MSE and MAPE are known as loss functions in determining the values of model weights[46]. In this paper, a larger class of loss functions is introduced. This class includes some well-known information measures. Suppose F and H are two distribution functions. General information is defined as follows:

$$GI = \frac{1}{G''(1)} \int G\left\{\frac{dF}{dH}\right\} dH, \tag{1}$$

where $\frac{dF}{dG}$ is the Radon-Nikodym derivative, $G(.)$is a univariate smooth function and $G''(1) \neq 0$ is the second derivative of $G(.)$at 1. Moreover, GI is well-defined if$GI(1) = 0$ and the first derivative of G is smooth at 1. Thefirst property makes $GI = 0$ if and only if$F = G$.The second property ensures the asymptotic statistical behavior, which is$2nGI_n \to \chi_v^2$ as $n \to \infty$, where, GI_n denotes GI when instead of the distribution function, the empirical distribution function is used [47]. If$G(u) = u\log(u)$, the famous Kullback-Leibler measure is obtained. It can be easily shown that the well-known information measures such as Pearson information, Triangular, Lin-Wang, etc. are a special case of Formula 1. Because of the well-defined general information, these measurements have the property of measuring the distance between two distributions, or even two vectors in a discrete state. In this paper, we use a number of these information measures including Kullback-Leibler, triangular and Lin-Wang as a loss function between two vectors. These information measures between discrete probability distributions$Q = (q_1, q_2, \dots q_n)andP = (p_1, p_2, \dots p_n)$ are introduced below[48].

a) Kullback-Leibler information measure

$$D_{KL}(P,Q) = \sum p_i \log\frac{p_i}{q_i}\cdot \tag{2}$$

b) Triangular information measure

$$D_T(P,Q) = \sum \frac{(p_i - q_i)^2}{p_i + q_i}. \tag{3}$$

c) Lin-Wang information measure

$$D_{LW}(P,Q) = \sum p_i \log\frac{2p_i}{p_i + q_i}. \tag{4}$$

[2]Weights are learnable parameters of neural networks which control the strength of the connection between two neurons. In other words, weights determine how much input can affect output.

Now, by substituting the target and output values with the probability distributions p and Q, these measures can be considered as a loss function. Next, we predict the price of bitcoin based on networks using these loss functions and compare it with commonly used loss functions such as MSE, MAE, and MAPE.

5 Experiment

5.1 Datasets and Methodology

In this paper, we used the bitcoin price data (USD), which can be provided easily from the link https://bitcoincharts.com/charts. The prices started daily from 10/01/2013 to 9/27/2020 and our goal is to predict them based on historical data. Assuming we have the data until the *i*-th day, our goal is to predict the price of the *(i+m)*-th day using the data sequence of [*i*-*k* to *i*], and we do this for different values of m including 10, 90, and 180.After pre-processing the data, including normalization and minimization, the training process begins. We use 70% of the data as a training set and the rest as a test set. In fact, we repeat the experiment ten times for the training set, and examine the model for the test set that provides the most accuracy for the training data and record it. The proposed method for predicting the bitcoin price is the regression capsule network (CapsNet) For the input data, the previous 225 data (*k*=225) is used, in the way that we transform the 225 by 1 input vector into a 15 by 15 matrix and use it as an input. The next layer is an ordinary convolution layer with 256 kernel number, 6 kernel size, and 1 stride. Then, the primary caps layer, where the features are divided into capsules, contains 32 kernels of size 6 and stride 2. After that, the operation squashing is performed. The outputs are then multiplied by the weight matrix to constitute the digit caps layer, where the squashing and dynamic routing operations are applied, and in the last step, after reshaping, a dense layer with 1 number of neuron is applied in order to predict the price.

5.2 Evaluation Metrics

Although in a problem such as regression, the measure of accuracy and goodness of fit is the average value of a function of predicted and the actual value, because one of our goals is to compare different loss functions with each other, we use some criteria which are applicable in practice for investors. These criteria, which are defined as follows, are commonly-used in papers related to machine learning, such as[49].

$$\text{Accuracy} = \frac{\text{UT} + \text{DT}}{\text{UT} + \text{UF} + \text{DT} + \text{DF}}. \tag{5}$$

$$\text{Precision1} = \frac{\text{UT}}{\text{UT} + \text{UF}}. \tag{6}$$

$$\text{Precision2} = \frac{\text{DT}}{\text{DT} + \text{DF}}. \tag{7}$$

Where UT is the number of times that the price increase, is correctly predicted.

DT, is the number of times that the price decrease, is correctly predicted.

UF, is the number of times that the price increase, is incorrectly predicted.

DF, the number of times that the price decrease, is incorrectly predicted.

Table 1: Comparison of the Accuracy for the Predicted Price of the Next 10 Days

	Criteria	Precision1		Precision2		Accuracy		Average
Net	Loss	Train	Test	Train	Test	Train	Test	
Caps Net	Lin	0.63	0.61	0.65	0.58	0.63	0.61	0.62
	KL	0.60	0.60	0.68	0.60	0.61	0.60	0.62
	Mse	0.64	0.65	0.50	0.47	0.55	0.54	0.56
	Mape	0.61	0.62	0.56	0.52	0.59	0.59	0.58
	Tri	**0.77**	**0.72**	**0.57**	**0.57**	**0.62**	**0.61**	**0.64**
ANN	Lin	0.59	0.56	0.52	0.53	0.57	0.56	0.56
	KL	0.57	0.56	0.53	0.62	0.57	0.56	0.57
	Mse	0.56	0.55	0.52	0.24	0.56	0.54	0.50
	Mape	0.54	0.56	0.43	0.51	0.50	0.55	0.52
	Tri	0.57	0.55	0.46	0.45	0.52	0.53	0.51
CNN	Lin	0.60	0.55	0.48	0.45	0.54	0.52	0.52
	KL	0.61	0.59	0.58	0.74	0.60	0.61	0.62
	Mse	0.63	0.60	0.53	0.61	0.58	0.60	0.59
	Mape	0.55	0.55	0.43	0.45	0.53	0.54	0.51
	Tri	0.66	0.58	0.57	0.56	0.62	0.58	0.60
LSTM	Lin	0.59	0.55	0.58	0.49	0.59	0.55	0.56
	KL	0.60	0.60	0.48	0.49	0.53	0.55	0.54
	Mse	0.62	0.62	0.48	0.50	0.53	0.56	0.55
	Mape	0.55	0.55	0.43	0.43	0.52	0.53	0.50
	Tri	0.58	0.54	0.56	0.41	0.58	0.52	0.53

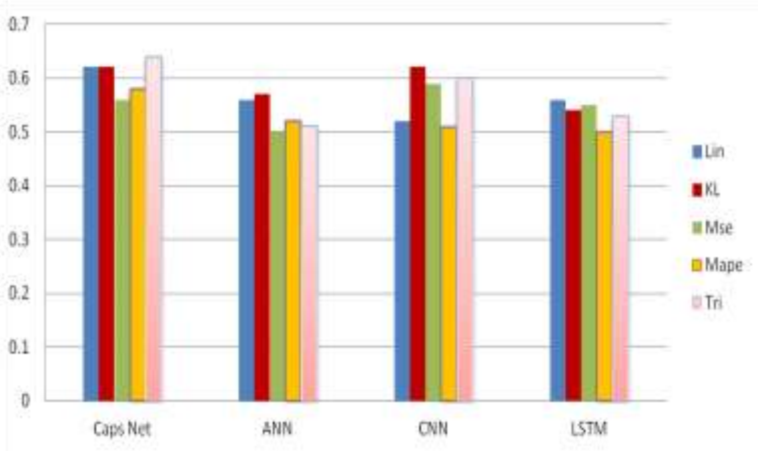

Fig. 2: Comparison of Accuracy for the Predicted Price of the Next 10 Days

5.3 Comparison Methods and Analysis

For the purpose of evaluating the goodness of the proposed model, we perform an experiment with several powerful and frequently-used prediction models. These techniques include ANN, CNN and

LSTM. Each of them uses different five loss functions, including MSE, MAPE, Kullback-Leibler information measure (KL), Lin-Wang information measure (LW) and Triangular information measure (T). Tables 1 to 3 and Figures 2 to 4 show the Accuracy of predicted bitcoin price values based on different models as well as different loss functions. As can be seen, criteria with formula 5, 6 and 7 have been used. The goodness of predictability was also measured for both training data and test data. As mentioned in the previous sections, we performed the experiment ten times on the training data, and the

Table 2: Comparison of the Accuracy for the Predicted Price of the Next 90 Days

	Criteria	Precision1		Precision2		Accuracy		Average
Net	Loss	Train	Test	Train	Test	Train	Test	
Caps Net	Lin	0.69	0.98	0.73	0.71	0.70	0.69	0.75
	KL	0.70	0.70	0.70	0.69	0.70	0.69	0.70
	Mse	0.71	0.69	0.68	0.63	0.70	0.67	0.68
	Mape	0.62	0.63	0.43	0.43	0.46	0.45	0.50
	Tri	**0.88**	**0.89**	**0.92**	**0.88**	**0.89**	**0.89**	**0.89**
ANN	Lin	0.67	0.69	0.95	0.92	0.72	0.73	0.78
	KL	0.66	0.69	0.74	0.79	0.69	0.71	0.71
	Mse	0.68	0.71	0.75	0.70	0.69	0.70	0.71
	Mape	0.64	0.66	0.97	0.93	0.68	0.70	0.76
	Tri	0.74	0.76	0.75	0.76	0.74	0.76	0.75
CNN	Lin	0.70	0.70	0.71	0.82	0.71	0.73	0.73
	KL	0.63	0.64	0.77	0.94	0.65	0.66	0.72
	Mse	0.74	0.69	0.83	0.82	0.77	0.72	0.76
	Mape	0.65	0.64	0.74	0.95	0.67	0.67	0.72
	Tri	0.86	0.75	0.68	0.77	0.77	0.75	0.76
LSTM	Lin	0.65	0.66	0.69	0.60	0.66	0.64	0.65
	KL	0.63	0.65	0.91	0.68	0.66	0.66	0.70
	Mse	0.68	0.70	0.52	.050	0.59	0.59	0.52
	Mape	0.64	0.66	0.86	0.72	0.67	0.67	0.70
	Tri	0.67	0.72	0.71	0.69	0.68	0.71	0.70

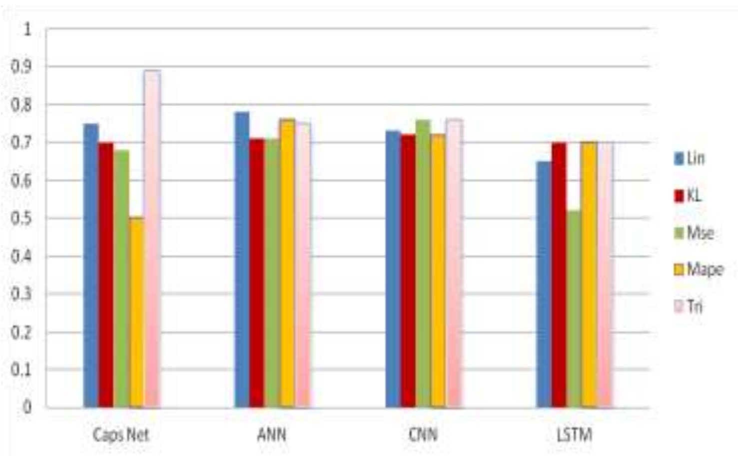

Fig. 3: Comparison of the Accuracy for the Predicted Price of the Next 90 Days

Table 3: Comparison of the Accuracy for the Predicted Price of the Next 180 Days

	Criteria	Precision1		Precision2		Accuracy		Average
Net	Loss	Train	Test	Train	Test	Train	Test	
Caps Net	Lin	0.85	0.86	0.80	0.80	0.84	0.84	0.83
	KL	0.72	0.73	0.98	0.96	0.76	0.76	0.82
	Mse	0.95	0.95	0.80	0.76	0.88	0.86	0.87
	Mape	0.61	0.64	0.00	0.00	0.61	0.64	0.42
	Tri	**0.95**	**0.96**	**0.95**	**0.89**	**0.95**	**0.93**	**0.94**
ANN	Lin	0.72	0.73	0.91	0.89	0.75	0.76	0.79
	KL	0.72	0.73	0.90	0.88	0.75	0.75	0.79
	Mse	0.71	0.71	0.80	0.76	0.73	0.71	0.74
	Mape	0.72	0.72	0.95	0.93	0.75	0.75	0.80
	Tri	0.74	0.75	0.80	0.77	0.75	0.76	0.76
CNN	Lin	0.82	0.85	0.74	0.80	0.79	0.83	0.81
	KL	0.70	0.71	0.99	0.96	0.74	0.74	0.81
	Mse	0.90	0.92	0.60	0.58	0.74	0.72	0.74
	Mape	0.68	0.68	0.98	0.94	0.71	0.70	0.78
	Tri	0.89	0.95	0.80	0.76	0.85	0.86	0.85
LSTM	Lin	0.73	0.70	0.77	0.55	0.73	0.65	0.69
	KL	0.72	0.70	0.73	0.55	0.72	0.65	0.68
	Mse	0.72	0.81	0.52	0.48	0.64	0.61	0.63
	Mape	0.77	0.75	0.53	0.50	0.65	0.63	0.64
	Tri	0.73	0.74	0.80	0.74	0.74	0.74	0.75

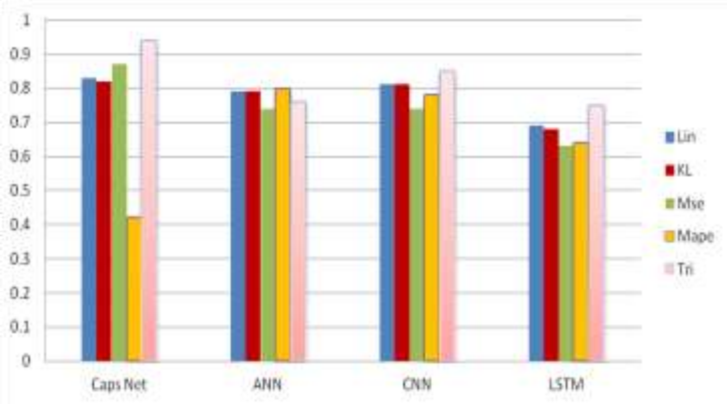

Fig. 4: Comparison of the Accuracy for the Predicted Price of the Next 180 Days

weights that provided the most accuracy on the training set were applied to the test set for comparison. The last column in the tables, average, shows the average of the forecast criteria. The bold row in the tables indicates the row that has achieved the highest average of accuracy. Table 1 shows the predicted price for the next 10 days. As can be seen, the capsule net with the triangular loss function has the highest accuracy among all other models. In the test data, 72% of the increased values, 57% of the decreased values, and in general, 61% of the directions are correctly predicted. Given that the above three criteria are calculated for both test data and training data (there are a total of 6 criteria for each network), the comparison is a bit difficult. Because each criterion is applicable according to each problem and condition. Therefore, we use their average to compare different networks. The average prediction criteria, for both the test and training sets, is 64%, which is higher than other models. Figure 2 also confirms this conclusion. Table 2 and Figure 3 indicates the predicted price for the next 90 days. Here, too, the greatest accuracy is related to the regression capsule with triangular loss, which is an average of89 percent correct in predicting the direction of the price. Finally, Table 3 and Figure 4 show the long-term forecast for the next 180 days, and it is clear that an average of 94% of the data are correctly predicted whether they will increase or decrease.

6 Conclusion

Forecasting financial markets, including the bitcoin, has always been of interest to researchers and investors. Because machine learning techniques, and especially deep learning, have had some success in this area, the use of tools such as ANN, CNN, and LSTM in prediction is very common. With the introduction of the capsule network in recent years and its remarkable success on MNIST data, this network is being developed on various datasets. Therefore, we also used this network to predict the price of bitcoin, with the change that we fist converted the classification network to a regression network and then used different information measures as a loss function. We made predictions for both the long-term and medium-term, and compared the results with powerful networks, which also use different loss functions, including MAE, MAPE, Kullback-Leiblerinformation measure (KL), Lin-Wang information measure (LW) and Triangular information measure (T). To compare the goodness of prediction, we used criteria that are applicable to investors in practice, such as accuracy and precision (Formula 5 to 7). These criteria are common for comparison of prediction and classification methods[49].Comparing the results, it is concluded that the regression capsule network with the triangular loss function has the best performance, especially for long-term prediction,so that the correct prediction percentages for the test and training data are 62 and 61, respectively, for the next 10 days, 89 and 89 for the next 90 days, and 95 and 93 for the next 180 days.On average, considering the total accuracy criteria for both the training and test data, in 64% of cases the prediction is correct in predicting the next 10 days. This average for the next 90 days is89%, and for the next 180 daysis 94%, which is higher than the other competitors.

Acknowledgment

The authors would like to thank the editor and anonymous reviewers for their constructive comments which led to the improvement of the paper.

Reference

[1] Kumar, D., Meghwani, S.S., Thakur, M., Proximal support vector machine based hybrid prediction models for trend forecasting in financial markets, Journal of Computational Science, 2016, 17, P.1-13.
Doi: 10.1016/j.jocs.2016.07.006.

[2] Podsiadlo, M., Rybinski, M., Financial time series forecasting using rough sets with time-weighted rule voting, Expert Systems with Applications, 2016, 66, P.219-233.

[3] Chen, Y., and Hao, Y., A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction, Expert Systems with Applications, 2017, 80, P. 340-355.
Doi: 10.1016/j.eswa.2017.02.044.

[4] Zhong, X., D. Enke, Forecasting daily stock market return using dimensionality reduction. Expert Systems with Applications, 2017,67,P. 126-139, Doi: 10.1016/j.eswa.2016.09.027.

[5] Cavalcante, R.C., Brasileiro, R. C., Souza, V. L.F., Nobrega, J.P., Oliveira, A. L.I., Computational intelligence and financial markets: A survey and future directions, Expert Systems with Applications, 2016, 55, P.194-211, Doi: 10.1016/j.eswa.2016.02.006.

[6] Pan, Y., Xiao, Z., Wang, X., Yang, D., A multiple support vector machine approach to stock index forecasting with mixed frequency sampling, Knowledge-Based Systems, 2017, 122, P.90-102.
Doi: 10.1016/j.knosys.2017.01.033.

[7] Zhang, N., Lin, A., Shang, P., Multidimensional k-nearest neighbor model based on EEMD for financial time series forecasting, Physica A: Statistical Mechanics and its Applications, 2017, 477, P. 161-173.
Doi: 10.1016/j.physa.2017.02.072.

[8] Le, H.H., Viviani, J., Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios, Research in International Business and Finance, 2018, 44, P.16-25.

[9] Lv, D., Huang, Z., Li, M., Xiang, Y., Selection of the optimal trading model for stock investment in different industries, PloS one, 2019, 14(2), P.e0212137. Doi: 10.1371/journal.pone.0212137.

[10] Chen, Z., Li, C., Sun, W., Bitcoin price prediction using machine learning: An approach to sample dimension engineering, Journal of Computational and Applied Mathematics, 2020, 365, P.112395.

[11] Olivas, E.S., Guerrero, J.D.M., Sober , M. M., Benedito , J, R, M., Lopez, A, J, S., Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques. 2009: IGI Global.

[12] McNally, S., Roche, J., Caton. S., Predicting the price of bitcoin using machine learning, 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing, 2018.
Doi: 10.1109/PDP2018.2018.00060

[13] Saad, M., M. Saad, J. Choi, D. Nyang, J. Kim and A. Mohaisen,., Toward characterizing blockchain-based cryptocurrencies for highly accurate predictions, IEEE Systems Journal, 2019, 14(1), P.321-332.
Doi: 10.1109/JSYST.2019.2927707.

[14] Jang, H. and J. Lee, An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information, IEEE Access, 2017,6, P.5427-5437.
Doi: 10.1109/ACCESS.2017.2779181.

[15] Rebane, J., Isak, K., Papapetrou, P., Seq2seq rnns and arima models for cryptocurrency prediction: A comparative study, SIGKDD Fintech, 2018,18,P. 2-6. Doi: 10.475/123_4.

[16] Shintate, T., Pichl, L., Trend prediction classification for high frequency bitcoin time series with deep learning, Journal of Risk and Financial Management, 2019,12(1),P.17. Doi: 10.3390/jrfm12010017.

[17] Davoodi Kasbi, A., Dadashi, I., Stock price prediction using the Chaid rule-based algorithm and particle swarm optimization (pso), Advances in Mathematical Finance and Applications, 2020,5(2),P.197-213.
Doi: 10.22034/amfa.2019.585043.1184.

[18] Ghasemzadeha, M., Mohammad-Karimi, N., Ansari-Samani, H., Machine learning algorithms for time series in financial markets, Advances in Mathematical Finance and Applications, 2020,5(4), P.479-490.
Doi: 10.22034/amfa.2020.674946.

[19] Hsu, M.-W., Lessmann, S., Sung, M.C., Ma, T., Johnson, J. E.V., Bridging the divide in financial market forecasting: machine learners vs. financial economists, Expert Systems with Applications, 2016,61,P.215-234.
Doi: 10.1016/j.eswa.2016.05.033.

[20] Patel, J., Shah, S.,Thakkar, P.,Kotecha, K., Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques, Expert systems with applications, 2015,42(1),P.259-268.
Doi: 10.1016/j.eswa.2014.07.040.

[21] Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., Soman, K. P., Stock price prediction using LSTM, RNN and CNN-sliding window model, International conference on advances in computing, communications and informatics, 2017.
Doi: 10.1109/ICACCI.2017.8126078.

[22] Sachdeva, A., Jethwani, G., Manjunath, C., Balamurugan, M., Krishna, A.V.N., An Effective Time Series Analysis for Equity Market Prediction Using Deep Learning Model, International Conference on Data Science and Communication (IconDSC), 2019.
Doi: 10.1109/IconDSC.2019.8817035.

[23] Liu, J., Lin, H., Liu, X., Xu, B., Ren, Y., Diao, Y., Yang, L., Transformer-based capsule network for stock movement prediction. in Proceedings of the First Workshop on Financial Technology and Natural Language Processing. 2019.

[24] Goodfellow, I., et al., Deep learning. Vol. 1. 2016,MIT press Cambridge.

[25] Fischer, T., C. Krauss, and A. Treichel, Machine learning for time series forecasting-a simulation study. 2018, FAU Discussion Papers in Economics.

[26] Korczak, J., Hemes. M., Deep learning for financial time series forecasting in A-Trader system. Federated Conference on Computer Science and Information Systems (FedCSIS). 2017.

[27] Zhou, X., Pan, Z., Hu, G., Tang, S., and Zhao, C., Stock market prediction on high-frequency data using generative adversarial nets, Mathematical Problems in Engineering, 2018,(1),P.1-11, Doi: 10.1155/2018/4907423.

[28] Fischer, T., Krauss, C., Deep learning with long short-term memory networks for financial market predictions, European Journal of Operational Research, 2018,270(2),P. 654-669, Doi:10.1016/j.ejor.2017.11.054.

[29] Gonçalves, R., Ribeiro, V. M., Pereira, F. L.,and Rocha, A. P., Deep learning in exchange markets. Information Economics and Policy, 2019. 47, P.38-51, Doi: 10.1016/j.infoecopol.2019.05.002.

[30] Tavakoli, M., Doosti, H., Forecasting the Tehran Stock market by Machine Learning Methods using a New Loss Function. Advances in Mathematical Finance and Applications, 2021,6(2),P.1-24.
Doi: 10.22034/amfa.2020.1896273.1399.

[31] Long, J., et al., An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market, Applied Soft Computing, 2020,P.106205.
Doi: 10.1016/j.asoc.2020.106205.

[32] Kelotra, A., Pandey, P., Stock market prediction using optimized deep-convlstm model. Big Data, 2020, 8(1),P.5-24. Doi: 10.1089/big.2018.0143.

[33] Bouktif, S., Fiaz, A., Awad, M., Augmented Textual Features-Based Stock Market Prediction, IEEE, 2020,8, P.40269-40282. Doi: 10.1109/ACCESS.2020.2976725.

[34] Chung, H., Shin, K., Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction, Neural Computing and Applications, 2020, 32(12), P.7897-7914. Doi: 10.1007/s00521-019-04236-3.

[35] Cervelló-Royo, R., Guijarro, F., Forecasting stock market trend: A comparison of machine learning algorithms, Finance, Markets and Valuation, 2020, 6(1), P.37-49.

[36] De Souza, M.J.S., Almudhaf, F.W., Henriqu, B,M., Negredo, A,B,S., Ramos, D,G,F., Sobreiro, V. A., Kimura, H., Can artificial intelligence enhance the Bitcoin bonanza, The Journal of Finance and Data Science, 2019, 5(2), P.83-98. Doi: 10.1016/j.jfds.2019.01.002.

[37] Ji, S., Kim, J., Im, H., A comparative study of Bitcoin price prediction using deep learning, Mathematics, 2019, 7(10), P.898. Doi: 10.3390/math7100898.

[38] Huisu, J., Predicting bitcoin prices by using rolling window LSTM model, Data Science in FinTech (DSF), 2018.

[39] Kim, Y.B., Kim, J.G., Kim, W., Im, J, H., Kim, T.H., Kang, Sh.J., Chang Hun Kim., Predicting fluctuations in cryptocurrency transactions based on user comments and replies, PloS one, 2016, 11(8), P.e0161197. Doi: 10.1371/journal.pone.0161197.

[40] Li, T.R., T.R., Li, Chamrajnagar, A.S., Fong, X.R., Rizik, N. R., Fu. F.,, Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model, Frontiers in Physics, 2019, 7, P.98. Doi: 10.3389/fphy.2019.00098.

[41] Banerjee, A., Merugu, S., Dhillon, I. S., Ghosh, J., Clustering with Bregman divergences, Journal of machine learning research, 2005, 6(Oct), P.1705-1749. Doi: 10.1137/1.9781611972740.22.

[42] Villmann, T., Haase, S., Mathematical aspects of divergence based vector quantization using Fréchet-derivatives, University of Applied Sciences Mittweida, 2010.

[43] Villmann, T., Haase, S., Schleif, F.M., Hammer, B., *Divergence based online learning in vector quantization*, *International Conference on Artificial Intelligence and Soft Computing*, 2010, Springer.

[44] Cichocki, A., Zdunek, R., Choi, S., Plemmons, R., Amari, Sh,-i.,*Novel multi-layer non-negative tensor factorization with sparsity constraints,* International Conference on Adaptive and Natural Computing Algorithms, 2007, Springer.

[45] Sabour, S., Frosst, N., Hinton. G.E., *Dynamic routing between capsules*, Advances in neural information processing systems, 2017.

[46] Gandhmal, D.P., Kumar, K., *Systematic analysis and review of stock market prediction techniques*, Computer Science Review, 2019, **34**, P.100190. Doi: 10.1016/j.cosrev.2019.08.001.

[47] Tilouche, S., Nia, V.P., Bassetto, S., *Parallel coordinate order for high-dimensional data*, arXiv preprint arXiv,1905.10035, 2019.

[48] Tavakoli, M., Alizadeh Noughabi, H., Mohtashami, G.R., *An estimation of Phi divergence and its application in testing normality,* Hacettepe Journal of Mathematics and Statistics, 2021, **6**(2), P.195-205. Doi: 10.15672/hujms.629192.

[49] Powers, D.M., *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.* arXiv preprint arXiv:2010.16061, 2020.