Electrical, Electronics and communications, and Computer Engineering

Impact of Twitter Sentiment Related to Bitcoin on Stock Price Returns

Feda Hassan Jahjah *
M.Sc. Software Engineering and Information Systems
Faculty of Informatics Engineering - Al-Baath University
Syria- Homs
Fjahjah@albaath-univ.edu.sy

Dr. Muhanad Rajab
Professor
Faculty of Informatics Engineering - Al-Baath University
Syria- Homs
Mrajab@albaath-univ.edu.sy

ABSTRACT

Twitter is becoming an increasingly popular platform used by financial analysts to monitor and forecast financial markets. In this paper we investigate the impact of the sentiments expressed in Twitter on the subsequent market movement, specifically the bitcoin exchange rate. This study is divided into two phases, the first phase is sentiment analysis, and the second phase is correlation and regression. We analyzed tweets associated with the Bitcoin in order to determine if the user's sentiment contained within those tweets reflects the exchange rate of the currency. The sentiment of users over a 2-month period is classified as having a positive or negative sentiment of the digital currency using the proposed CNN-LSTM deep learning model. By applying Pearson's correlation, we found that the sentiment of the day (d) had a positive effect on the future Bitcoin returns on the next day (d+1). The prediction accuracy of the linear regression model for the next day's revenue was 78%.

Keywords: Sentiment Analysis, Deep Learning, Mining Social Media, Correlation, Multiple Linear Regression.

تأثير المشاعر المتعلقة ببيتكوين على تويتر على عائدات سعر السهم

فدا حسن جهجاه
ماجستير هندسة البرمجيات ونظم المعلومات
كلية الهندسة المعلوماتية- جامعة البعث

د. مهند رجب
أستاذ
كلية الهندسة المعلوماتية- جامعة البعث

الخلاصة

أصبحت وسائل التواصل الاجتماعي عامةً وخاصة خدمة التدوين المصغر تويتر Twitter، منصة شعبية للأفراد، ومنهم المحللون الماليون، لتبادل آرائهم وتحليلاتهم حول الأسواق المالية. نقوم في هذا البحث بدراسة تأثير المشاعر المعبر عنها عبر تويتر على حركة السوق اللاحقة، وتحديداً سعر صرف عملة البيتكوين. حيث تنقسم هذه الدراسة إلى مرحلتين: تحليل المشاعر،

والارتباط والانحدار. تم تحليل التغريدات المرتبطة بالعملة الرقمية بيتكوين من أجل تحديد ما إذا كان شعور المستخدم المتضمن في هذه التغريدات يعكس سعر صرف العملة. حيث تم تصنيف شعور المستخدمين على مدى فترة شهرين تقريباً تجاه العملة الرقمية على أنها تحتوي شعوراً إيجابياً أو سلبياً باستخدام نموذج التعلّم العميق المقترح CNN-LSTM. وبتطبيق ارتباط بيرسون وجدنا أن درجة الشعور في اليوم لها تأثير إيجابي على عوائد البيتكوين المستقبلية في اليوم الذي يليه. وحقق نموذج الانحدار الخطي للتنبؤ بعائدات اليوم التالي دقة 78%.

الكلمات الرئيسية: تحليل المشاعر، التعلّم العميق، التنقيب في وسائط التواصل الاجتماعية، الارتباط، الانحدار الخطي المتعدد.

1. INTRODUCTION

Sentiment analysis is generally used to identify the writer's feelings about a topic that can express an opinion or emotional state. It aims to classify the polarity of a text according to the opinion of the writer, by detecting positive, negative or neutral sentiments about a particular subject. It is used in marketing, customer service and other fields**. (Taboada et al., 2011)**. In particular, companies are interested in knowing their customers' opinions about their products and services, they are making a lot of effort and money to do so. Considering that the Twitter includes millions of people exchanging views and feelings with the hypothesis of a relationship between public opinion and real-world events. Thus, this paper will deal with small texts from Twitter.

More recently, deep-learning applications have shown remarkable results in natural language processing including sentiment analysis across multiple data sets **(Collobert et al., 2001)**. These models don't need to be provided with predefined features, but they learn advanced features from the data set by themselves. The words are represented in a high-dimensional vector space, and the neural network does the feature extraction process. Structures such as RNN are able to understand sentence structure efficiently. These characteristics make deep learning models naturally suited to the task of sentiment analysis.

This paper aims to analyze the sentiment of Twitter data related to Bitcoin, which can serve as a predictive basis for currency returns. Therefore, this work relies on two main concepts: sentiment analysis and correlation. To implement sentiment analysis using suggested deep learning model, we collected tweets for five months of 2018, a period that also allows validation of results with a high degree of confidence. For correlation we collected stock data for two months. We'll describe the methodology used, methods related to collecting Twitter data, sentiment analysis, associate sentiment with stocks, and accurately explain the results.

2. PREVIOUS WORK

Analyzing and predicting the behavior of stocks in the stock market has attracted much attention from academics as well as investors. A model that takes into account investors' sentiment and interests can provide a better explanation of stock market returns, and can also contain useful information to predict stock market returns.

(Bollen et al., 2011) analyzed public tweets from February 28 to December 19, 2008. Tweets were filtered using some general emotional expressions (such as I'm feeling). They analyzed the text content by two mood-tracking tools based on general linguistic dictionaries: Opinion Finder that measures positive and negative mood, and Google-Profile of Mood States (GPOMS) that measures mood in terms of 6 dimensions (Calm, Alert, Sure, Vital, Kind, and Happy). A Granger causality analysis and Self-Organizing Fuzzy Neural Network were used to prove the hypothesis that once public opinion is known, changes in DJIA (Dow Jones Industrial Average) closing values can be predicted with an accuracy of 87.56%.

(Xu et al., 2012) suggested a way to analyze the public sentiment and Web mining to predict financial markets for the S & P500 Index. In the proposed method they used a static data set (SNAP), which constitutes 20-30% of the public tweets that published from June 1 to December 31, 2009, contains 467 million texts from 20 million users. They displayed the emotions embedded in the tweet using Opinion Finder to categorize 5 classes (strong positive, weak positive, normal, strong negative, weak negative). The SVR model was used to model the relationship between online sentiment and financial market prices. It proved that the selective model is good for predicting the financial market. The study showed that using the function (STEF) to integrate daily emotions there was no obvious improvement on the prediction performance.

(Stingfest and Luno, 2017) used the Bayesian method to analyze the relationship of 2.27 million tweets over a 31-day period linked to Bitcoin with a change in the stock price in the near future based on periods ranging from 5 minutes to 4 hours and accurately 79%. They used the VADER tool (Valence Aware Dictionary and sEntiment Reasoner) to detect the polarity (negative - neutral - positive) and intensity of sentiment in the text.

(Colianni et al., 2015) analyzed how Twitter's feelings can be used to influence investment decisions specifically on Bitcoin. The authors used the supervised machine learning techniques SVM (Support Vector Machines), logistic regression and Bayesian algorithm that resulted an accuracy of over 90% per hour throughout the day.

(Ranco et al., 2015) they used Twitter data for 30 DJIA companies over a 15-month period, and found Pearson correlation and Garner's causality relatively low between the time series over the entire time period studied. A static dataset and SVM classification model were used taking into account the total number of tweets per day, the number of positive tweets per day, the number of negative tweets per day, and the polarity of positive or negative sentiment.

(Sohangir et al., 2018) they applied several deep learning models such as Long-Short Term Memory LSTM, doc2vec and Convolutional neural networks to stock market opinions published in StockTwits, the results show that the deep learning model can be used effectively to analyze financial sentiments and that the Convolutional neural network is the best model To predict the sentiment of the authors in the StockTwits dataset.

(Kar et al., 2017) also introduced a sentiment analysis system to extract sentiment from microblogging and news headlines. Lexical and emotional vocabulary features and metadata were used with representations derived from Convolutional neural networks (CNN) and a bi-directional gateway replication unit (Bi-GRU). Two systems are designed to predict the value of sentiment polarity. The first system takes advantage of hand-designed features and uses Vector Regression (SVR) to predict emotion outcomes. The second system combines engineering features with representation learned using CNN and Bi-GRU to predict emotion outcomes.

In short, previous researches have shown that public sentiment is a real factor in influencing investors. In this paper, we will investigate the performance and effectiveness of the proposed deep learning model resulting from the integration of CNNs and LSTMs in improving the accuracy of sentiment analysis. And study the correlation of emotions extracted and their impact on Bitcoin currency returns as well as to verify their predictive ability of these returns using Bitcoin-related Twitter data.

3. BACKGROUND

In this section we will briefly elaborate on the concept of deep learning and the basic model constructs of Convolutional Neural Networks and Long-Short Term Memory Neural Networks.

3.1 Deep Learning

Deep learning is a special type or form of machine learning that includes algorithms and techniques that allow the machine to learn by simulating neurons in the human body .Studies in this field have shown significant progress and effectiveness in various fields including facial recognition, computer vision and natural language processing.

Models are trained using a wide range of pre-labeled data (class-specific) and using multi-layer neural network structures that learn properties directly from the data without the need for manual extraction of features **Fig. 1**. For this reason, deep learning models are known as deep learning networks. The word Deep Refers to the number of hidden layers. Traditional networks contain 2-3 layers while deep networks can contain tens and hundreds of hidden layers.
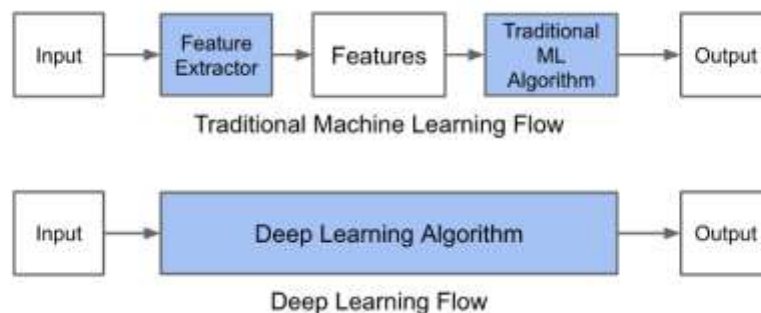

Figure 1. Deep learning model and traditional machine learning techniques.

3.2 Convolutional Neural Networks (CNN)

Convolutional neural networks is a special type of Feed forward neural network **(Jones, 2017) .** These networks are designed for image processing, classification issues ,video recognition and various tasks in natural language processing. CNNs have the ability of recognizing local features inside a multi-dimensional field. **Fig. 2** shows that there are 3 basic components of network identification:

1. Convolution Layers: composed of multiple filters that learn different features from multidimensional input data (e.g. images, word embedding, etc.). These filters are sequentially applied to different sections of the input.
2. Subsampling or pooling layers: a technique to compress or generalize feature representations and generally reduce the over fitting of the training data by the model (reduce dimensions of the extracted features) while retaining the most important information. Pooling layers are often very simple, taking the average or the maximum of the input.
3. Output Layer: Fully connected layer, a set of nodes to produce an output equal to the number of classes.
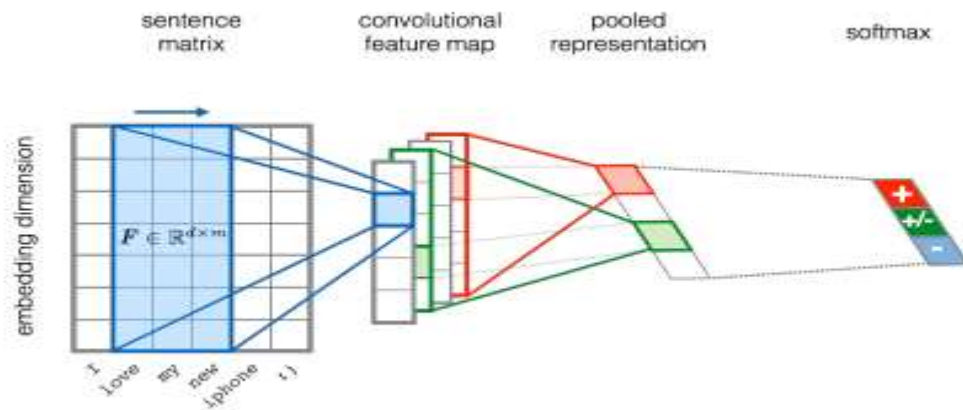
Figure 2. Convolutional Neural Networks for Sentence Classification.

3.3 Long-Short term memory Neural Network (LSTM)

Long-Short Term Memory (LSTM) network is a type of Recurrent Neural Network that is trained using Back propagation. It was first proposed by **(Hochreiter and Schimdhuber, 1997)**. Instead of neurons, LSTM networks have memory blocks that are connected into layers. A memory block can retain its value for a short or long period, allowing the cell to remember what is important and not just the last calculated value.

A block contains three gates that manage the block's state and output, and control how information flows into/out of the cell:

1. Input Gate: Responsible for adding new information and controls the state of the flow to memory.
2. Forget Gate: Responsible for deleting unimportant information from the cell. It produces a 0 or 1 output that is multiplied by the internal cell state variable.
3. Output Gate: Controls how much the value stored in memory affects the output activation of the block.

The cell also contains weights that control each gate. The training algorithm BPTT ,Will improve these weights based on the resulting network output error.

4. DATASETS

In this paper, two datasets are used in the experiments. The first is Twitter Tweets related to the Bitcoin with two sentiment classes, i.e., Positive and Negative and the second is the historical data of the Bitcoin exchange rate BTC / USD .

Twitter dataset: related to Bitcoin. From January 1 to May 31, 2018 tweets affecting the stock were collected which includes the symbols $ (cashtag) and # (hashtag) followed by the stock

symbol such as **$ btc,** #bitcoin. Using an open source Java software tool[1] which returns data as a csv file, each line consists of the following tweet information:

- Id: Twitter ID.
- Retweeted: Tweets retweet by other users.
- Favorites: The total number of tweet preferences.
- Username: The name of the user who posted the Tweet.
- Text: Tweet text.
- Mentions.
- Hashtags.
- Date: Tweet date in the format yyyy / mm / dd   HH: MM.

The dataset contains 25,746 tweets with two classes (Positive and Negative) and neutral has been neglected) and each class contains 5000 tweets. We use VADER tool: a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media *(***Hutto & Gilbert, 2014***).*

Historical Bitcoin exchange rate dataset: were collected from the site Coindesk. Depending on the (day) period as the length of time period for the requested data. We obtained OHCL data which represents the Open, High, Close, and Low values during the period studied from February 8 until March 31, 2018. The data has processed to become suitable for reliable analysis. The main problem was the availability of Twitter data for each day of the week, including weekends, and the absence of stock values on weekends and other holidays when the market is close. Using a simple statistical function, the missing values are filled. If the missing value is y ,the previous known value is xprevious ,the next known value is xnext ,and the value y will calculate by Eq. (1):

$$y = (xprevious + xnext)/2 \tag{1}$$

5. SENTIMENT ANALYSIS USING DEEP LEARNING MODEL

We will test the performance of the proposed deep learning model, which is a combination consists of an initial convolution network CNN and LSTMAs as in **Fig. 3**. The input data layer is represented as a matrix with a fixed dimension of vectors using GloVe Pre-trained (word embedding as input) . Multiple convolutional filters are applied to produce a new feature map . Then the Max-pooling layer calculates the maximum value of a feature corresponding to a particular filter (the most important features), its output will then fed into an LSTM layer.
LSTM layer measure the long-term dependencies of the sequence of features. Its output then fed into a fully connected layer where the sigmoid function is applied as an activation function. Finally the output layer generates the final output: either positive or negative.

In this model, the convolution layer extracts the local features to enable the layer LSTM Then use the order of these features to identify the order of the input text .
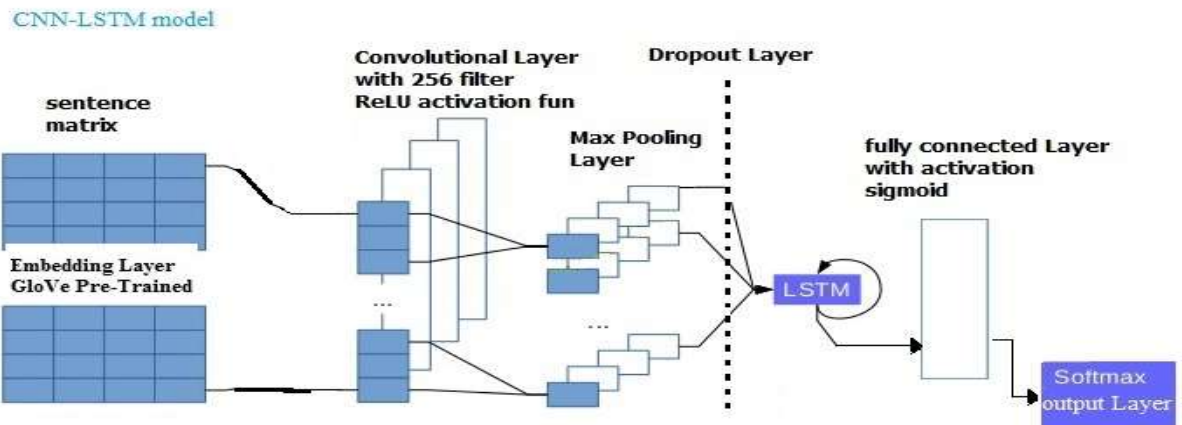
---

[1] https://github.com/Jefferson-Henrique/GetOldTweets -java/

Figure 3. Structure of the proposed CNN-LSTM sentiment analysis model.

5.1 Experiments and Results

A. About Keras

Keras is an open source neural network library written in Python. It defines high-level neural networks running on top of either TensorFlow[2] or Theano[3]. Keras contains numerous implementations of commonly used neural-network building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to make working with image and text data easier. It has support for convolutional and recurrent neural networks.

B. Parameters

The proposed model was built using keras. For this test 70% of Tweets (7000 tweets) were approved as a training group and 30% (3000 tweets) as a test group. These training and testing sets contained equal amount of positive and negative tweets. We trained the model using the parameters presented on **Table 1.** Obtained from the tests, and recorded the model's accuracy when trying to classify the test set.

Table 1. Parameters selected for our model

Parameter	Value
Embedding Dimension	300
Epoch	10
Batch Size	64
Filters	256
Kernel Size	3
Pool Size	2
Dropout	0.2
Word Embedding	Pre-trained (GloVe)

---

[2] https://github.com/tensorflow/tensorflow

[3] https://github.com/Theano/Theano

C. Result Analysis

The performance of sentiment classification can be evaluated by using the accuracy of the classifier Eq. (2):

$$accuracy = (TP + TN)/(TP + TN + FP + FN) \qquad (2)$$

In which, TP is the number of tweets that are positive and predicted correctly positive, TN is the number of tweets that are negative and predicted correctly as negative, FP is the number of tweets that are negative but predicted incorrectly as positive, and FN is the number of tweets that are positive but predicted incorrectly as negative.

It is clear from **Table 2.** That the proposed model improved the performance of sentiment classification and achieved higher score of about 1.8% - 9.5% better than other models. CNN-LSTM model achieved an accuracy of 4% higher than the CNN model ,and 1.9% better than the model LSTM. Combining CNNs and LSTMs Enabled us to use both capacity of CNN To recognize local patterns and choose good features, and ability LSTM to learn sequential data. Because the convolution layer provides important information, the layer LSTM succeed in harnessing full potential. Thus, the model achieves its full potential . The **Fig. 4** shows the increase accuracy and decrease the loss of the model over time (through 10 cycles).

Table 2. Compare the accuracy of the proposed method with other models for the same dataset.

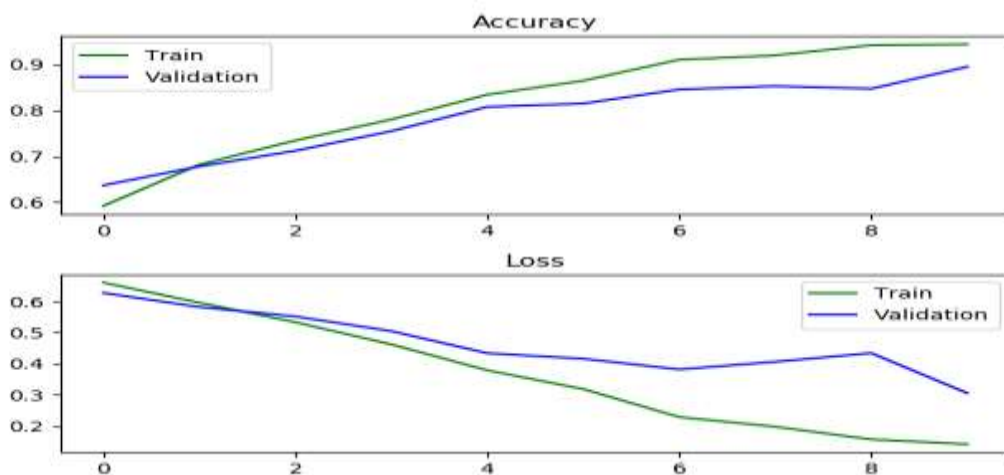| model	Average accuracy
CNN-LSTM	88.7%
CNN	84.6%
LSTM	86.83%
Naïve Bayes	79.2%
SVM	81.4%

Figure 4. Increasing accuracy and decreasing loss of the proposed model within 10 cycles.

6. IMPACT OF TWITTER SENTIMENT ON BITCOIN STOCK RETURNS

The goal is to study the correlation between sentiment and stock return. Hence we applied the Pearson correlation between the two variables then we build a prediction model using multiple linear regression.

6.1 Correlation

Using d Pearson correlation coefficient to determine the nature and strength of the linear relationship between two variables, where we have two variables: the first independent variable X, The second is a dependent variable Y its result is not specified and depends on the values of the independent variable **(Rajan and Nadu, 2016).** The correlation coefficient (r) is a value ranging between -1 and +1 that measures the correlation strength. A positive coefficient indicates a positive relationship, while a negative one indicates an inverse relationship. A value of zero means that there is no relationship. The following points are the accepted guidelines for interpreting the correlation coefficient:

- Values between 0 and 0.3(0 and -0.3) indicate a very weak positive (negative) linear relationship.
- Values between 0.3 and 0.5(-0.3 and -0.5) indicate a weak positive (negative) linear relationship
- Values between 0.5 and 0.7(-0.5 and -0.7) indicate a moderate positive (negative) linear relationship.
- Values between 0.7 and 1(-0.7 and -1) indicate a strong positive (negative) linear relationship.

The correlation coefficient is calculated by Eq. (3):

$$r_p = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}} \tag{3}$$

6.2 Multiple linear regression

In general purpose of multiple regression is to learn more about the relationship between several independent or predictor variables and a dependent variable. Multiple regression is a flexible method of data analysis used as a predictive model. The dependent variable can be estimated by the value of several independent variables using the following regression Eq. (4):

$$Y_i = \beta_0 + \sum \beta_i X_i + \varepsilon_i \tag{4}$$

Where the β_i is the slope of the regression line which represents the effect of X_i on Y, ε_i Random.

6.3 Variables

6.3.1 Independent Variable

We considered the distinct influence of each tweet, according to its characteristics, for each tweet, we considered the two variables:

- Retweeted: the number of times a tweet had been retweeted.
- Favorites: the number of times a tweet had been listed as a "favorite".

We classified all tweets from February 8 to March 31 using the proposed deep model CNN-LSTM then we chose the best 25 tweets per day according to the two previous variables and we consider them to be effective, because if the tweet has a positive rating but it does not have a good number

of retweet and favorites, it can't be considered an influential tweet in public opinion. The two variables will be taken to give strength to the tweet. The degree of sentiment is calculated for a day (d) According to the following Eq. (5):

$$sd = \frac{tw^+ - tw^-}{tw^+ + tw^-} \tag{5}$$

Where tw $^-$ is the number of Negative Tweets and tw $^+$ is the number of Positive Tweets.

6.3.2 Dependent Variable

To calculate the daily return of the Bitcoin stock price we took the difference between the closing price on the day (d) and the closing price on the previous day (d-1) According to the Eq. (6) :

$$Rd = \frac{close_d - close_{d-1}}{close_{d-1}} \tag{6}$$

6.4 Results and Discussion

The Pearson correlation coefficient was calculated using the function pearsonr (Y, X) in Python within the library scipy.stats.stats. To study the correlation between the degree of sentiment, the closing price ,and the stock returns for the next day (d + 1)and two days later (d + 2), and after three days (d + 3). **Table 3.** Represents the Pearson correlation coefficient between variables.

Table 3. Correlation between sentiment and returns.

	close $_d$	S $_d$
R $_{d+1}$	0.63	0.69
R $_{d+2}$	0.28	0.35
R $_{d+3}$	0.1	0.1

The relationship between the sentiment and the returns of the next day (d + 1) is moderate and positive because the 'r' value was 0.69 and validate the condition $0.5 \leq r < 0.7$. The correlation is considered as statistically significant because the value 'p-value' is smaller than 0.05.

The relationship between the sentiment and the returns of the next two days (d + 2) is weak and positive because the 'r' value was 0.35 and validate the condition $0 \leq r < 0.5$. The correlation between the returns after three days and sentiment was very weak ('r' value =0.1) and could not be adopted.

Therefore, depending on Pearson correlation coefficient, the stock returns on day is correlated with both the sentiment and closing price of the stock on the previous day.

In this study we used multiple regression analysis to determine whether the sentiment is a good predictor of its Stock returns. Here the dependent variable is stock returns on day while the independent variables are Closing price and Sentiment degree of a stock on the previous day. We get the following result:

- R- squared: 0.713 which reflect the relevance of the model. That is %70 of the changes in the dependent variable resulting from independent variables.
- Prob (F-statistic) = 5.42e-08 which meets the condition: smaller than 0.05 and hence an acceptable value.
- Durbin-Watson= 2.03 the value is an acceptable value (between 1.5 and 2.5).

Table 4. Significant Values for Independent Variables.

Independent variable	P> | t |	std err
close	0.019	0.12
sentiment	0.014	0.141

Table 4. Represents the value of significance (P> | t |) and error(std err) which reflects the level of accuracy of variables. Anything below 0.05(P<0.0.5) is considered to be statistically significant (the variable can be used as a predictor). In this case, both closing price (p=0.019) and the sentiment (p=0.014) are below 0.05, and therefore they are considered to be a good predictors of stock returns and have an impact on it. The model can be considered acceptable and the regression line equation is as follows:

$$StockReturn = (\text{const coef}) + (Close\ \text{coef}) * X1 + (\text{Sentiment coef}) * X2$$

Applying the test dataset, which represents 30% (1200 tweets) of the study data in this part, to the predictive model we obtained an accuracy of 78%. Therefore we can be said: that the Twitter sentiment related to Bitcoin linked to the stock returns on the next day and have a good predictive ability.

7. CONCLUSIONS AND FUTURE WORK

In this study we discussed the relationship between Twitter and Bitcoin performance. We have presented a model that aim to combine CNN and LSTM neural networks to achieve better performance on sentiment analysis tasks. The model achieved an accuracy rating of 88.7% .The proposed model was compared with reference algorithms and using the same data set for both machine learning and deep learning methods. Finally, the correlation between sentiment and stock returns was conducted by applying the Pearson correlation which showed a positive moderate correlation between variables. A prediction model was constructed using multiple linear regression and we obtained a prediction accuracy of 78% .

In the future work, it would be beneficial to test the impact of other factors on the strength of the tweet, such as the influence of the twitter user, the number of followers associated with it. Also more predictive techniques can be examined such as Real-time neural networks or deep learning networks, especially LSTM or try some nonlinear methods.

REFERNCES

- Bollen, J., Mao, H., & Zeng, X., 2011. Twitter mood predicts the stock market. *Journal of Computational Science, 2 (1)*, pp.1, 8.

- Colianni, S., Rosales, S., & Signorotti, M., 2015. Algorithmic trading of cryptocurrency based on twitter sentiment analysis. Available at: .http://cs229.stanford.edu/proj2015/029_report.pdf .

- Collobert, Ronan, et al. 2001. Natural language processing (almost) from scratch. *The Journal of Machine Learning Research, 12,* pp.2493, 2537.

- Hochreiter, S. & Schmidhuber, J., 1997. Long short-term memory. *Neural Computation, 9(8)*, p.1735, 1780.

- Hutto, CJ & Gilbert, EE, 2014. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Ann Arbor*, MI.

- Jones, MT, 2017. Deep learning architectures the rise of artificial intelligence. *developerWorks.*

- Kar, S., Maharjan, S., & Solorio, T., 2017. RiTUAL -UH at SemEval-2017 Task 5: Sentiment Analysis on Financial Data Using Neural Networks. *Proceedings of the 11th International Workshop on Semantic Evaluations (SemEval-2017)*, pp.877, 882. Vancouver, Canada, Association for Computational Linguistics.

- Lönnö, J. & Stenqvist, E., 2017. Predicting Bitcoin price fluctuation with Twitter sentiment analysis, STOCKHOLM, SWEDEN: KTH ROYAL *Institute of Technology school of computer science and communication*

- Rajan, P., and Nadu, T., 2016. Web Sentiment Analysis: Comparison of Sentiments with Stock Prices Using Correlation and Regression. *Journal of Computing Technologies, 3(5), pp.2278, 3814.*

- Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I., 2015. The Effects of Twitter Sentiment on Stock Price Returns. *PLoS ONE,* 10 (9).

- Sohangir, S., Wang, D., Pomeranets, A., & Khoshgoftaar, T., 2018. Big Data: Deep Learning for financial sentiment analysis. *Journal Big Data, 5(3).* Available at: <https://doi.org/10.1186/s40537-017-0111-6.

- Taboada M., Brooke J., Tofiloski M., Voll K., and Stede M., 2011. Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics, 37(2)*, pp. 267, 307.

- Xu, W., Li, T., Jiang, B., & Cheng, C., 2012. Web Mining for Financial Market Prediction Based On Online Sentiments. *PACIS 2012 Proceedings*.