ISSN 2278-3091
Volume 9, No.4, July – August 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse351942020.pdf
https://doi.org/10.30534/ijatcse/2020/351942020

Predicting Crypto Currency Prices Using Machine Learning and Deep Learning Techniques

Lokesh Vaddi[1], Vaishnavi Neelisetty[2], Bhavana Chowdary Vallabhaneni[3], Kolla Bhanu Prakash[4]
[1]Dept. of CSE, K L Deemed to be University, India,vaddiv411@gmail.com
[2]Dept. of CSE, K L Deemed to be University, India,neelisettyvaishnavi@gmail.com
[3]Dept. of CSE, K L Deemed to be University, India,vallabhanenib2123@gmail.com
[4]Professor & Research Group Head, Dept. Of CSE, K L Deemed to be University, India, drkbp@kluniversity.in

ABSTRACT

In the past eight years of Bitcoin's history, the economy has seen the price of Bitcoin rapidly grow due to its promising outlook on the future for crypto currencies. Investors have taken note of several advantages Bitcoin provides over the traditional banking system. One such trait is that Bitcoin allows for decentralized banking, meaning that Bitcoin cannot be regulated by powerful banks. There is also a market cap of 21 million Bitcoins that can be in circulation, therefore a surplus of Bitcoins cannot be "printed" which would result in inflation. Bitcoin resolves the issue with transaction security by using a block chain, or a ledger, which records the history of every transaction ever made into one long hexadecimal "chain" of anonymous transactions, which keeps transaction history transparent, but also confidential. Bitcoin as a result has become a very bullish investing opportunity, and due to the huge volatility of the Bitcoin market price, this paper attempts to aid in investment decision making by providing Bitcoin market price prediction. Our team explored several Machine Learning algorithms by using Supervised Learning to train a prediction model and provide informative analysis of future market prices. We start with Linear Regression models, and train on several important features, then We proceed with the implementation of Recurrent Neural Networks (RNN) with Long Short Term Memory (LSTM) cells. All code is written in Python using Google's Tensor Flow software library. We show that the price of Bitcoin can be predicted with Machine Learning with high degree of accuracy.

Key words: Bitcoins, Crypto Currency, Deep Learning, Machine Learning

1. INTRODUCTION

The Bitcoin crypto currency experienced tremendous growth over the past year. The price of one Bitcoin went from about $750 at the end of 2016 to over $10,000 in the mid of 2020. Great rates of growth were also observed in the other crypto currencies, such as Ethereum and Litecoin. There is currently a shortage of quantitative analysis tools and techniques for predicting the prices of crypto currencies. Mathematical analysis has a well-established place in the financial industry for evaluating expected returns of a stock of a given company or performance of an entire portfolio. [12] However, Machine Learning & Deep Learning literature is lacking verification of whether the stock analysis techniques are valid for the crypto currencies, and if so, how they can be modified. [13] That is, what features need to be removed or added as a basis for price prediction, whether current Machine Learning algorithms & Deep Learning work for crypto currencies, and which approach yields the best results. [14][15][16][17] In this paper, we investigate these questions. Such analysis is relevant given a great amount of attention that crypto currencies, in particular Bitcoin, are generating. [18] Both individuals and large financial firms are attracted to crypto currencies because of the transparency and anonymity that they provide to their users, as well as their resistance to fraud due to the distributed nature of the ledger records. [19] Moreover, purchasing crypto currencies is promising in terms of making a profit and should be of interest to investors. [20] In addition to familiarizing themselves with industry trends and political and economic news, they can utilize Machine Learning & Deep Learning models to help them decide to buy or sell crypto currency. As described in greater detail in the future sections of this paper, Machine Learning & Deep Learning algorithms can be applied to crypto currency data to predict future price movements. [21] Different approaches produce results of different accuracies. [22] However, given a strong correlation between the actual direction of the price change and prediction of the algorithms, we were able to show that mathematical analysis can aid investors' decision making considerably and allow one to increase the chances of making a profit by trading crypto currencies. Quantitative prediction of the prices of crypto currencies can also be utilized to build an autonomous agent that performs trades on behalf of the investor based on historical price information, current news, sentiment analysis, and real-time data. [23] Our work only evaluates Machine Learning & Deep Learning models trained on historical data, but it can be extended to building and trading financial securities. [24] In this paper, we provide a background overview of current prediction modeling techniques being used by stock exchanges and follow up with several differences as well as similarities that can be made by treating crypto currencies as one would a share of stock. We then proceed by stating how crypto currency prediction modeling is a relatively young technique that has not been fully explored in the research literature. [25][26] The methodology section contains a description of the approach we followed to ascertain whether the price of Bitcoin can be predicted with Machine Learning & Deep Learning. [27] In particular, we discuss the application of Linear Regression and Recurrent Neural Networks (RNNs) with Long Short Term Memory (LSTM) cells to the task of predicting the market price of Crypto currencies, particularly on Bitcoin. [28] We identify relevant features for price prediction and determine the accuracy of single- and multi-feature Machine Learning models. [29] In the results section, we show and evaluate the results of the experiments performed on the historical price of Bitcoin by comparing the actual crypto currency price in the past with the predicted values. In the past work section, we provided an overview of relevant topics in the research literature concerning using Machine Learning & Deep Learning in the financial industry. Finally, we present concluding remarks and suggest the direction for future work on the subject of applying Machine Learning & Deep Learning to crypto currency price prediction.

2. LITERATURE SURVEY

Due to the high volatility associated with the market price of Crypto currencies like Bitcoin, prediction models have become more and more challenging for investors to accurately forecast investment decisions. [1] Thinking in regards to a stock market investment, a typical approach many investors would use as a prediction metric, is called Fibonacci retracement. [2] Fibonacci retracement helps identify the opportune moments for traders to buy or sell stock at the optimal price. [3] Fibonacci retracement utilizes the volatility of the stock, meaning that the two extreme trends in a graph are found, and the calculated vertical distance is then divided by key Fibonacci ratios, most commonly the golden ratio of 0.618. [4] These values are then used to help identify the most prominent support and resistance levels.

The problem with such prediction models however, is that the Fibonacci retracement levels are static, and are constant calculated values that are only good metrics for simple identification. In regards to crypto currencies like Bitcoin, time series predictions are not a novel idea for brokerage companies, and many brokers have invested heavily to improve performance of forecasting stock prices and foreign exchange rates by way of Machine Learning & Deep Learning. Bitcoin, a crypto currency first launched in 2009 by Satoshi Nakamoto, has parallels with many of the stock markets such as the New York Stock Exchange (NYSE) and can be modelled in similar ways. [5]

However, there are key differences between crypto currencies like Bitcoin and the stock market which makes our prediction model much different in nature. For one, Bitcoin commands a much smaller transaction volume as compared to many of the major stock exchanges. The average daily turnover for Bitcoin at the present day is $11 Billion, whereas trillions are seen for the latter. As a result of the small transaction volume, lower liquidity, and huge peak in investment turnover of crypto currencies like Bitcoin as the future of decentralized banking increases, the market price of crypto currencies like Bitcoin has become highly volatile and as a result, challenging to make successful predictions. [6]

While there is a great amount of scholarly literature and practical guides on Machine Learning & Deep Learning applied to the stock market, there is a lack of information about predicting movement of the value of crypto currencies. Namely, there is no consensus on the features that are to be used for training a prediction model. Our research primarily focuses on block chain size as an important feature of the market price trend, using various Machine Learning & Deep Learning models, as well as expanding our input variables into multi featured data. Previous attempts at time series forecasting have been well-documented and there are a lot of resources and tools available to use for free which our team plans to utilize in order to solve the problem at hand efficiently rather than reinventing the wheel at each stage. Specifically, Quandl API is used as the primary distributor of large feature data, and Google's tensor Flow library for Python allows for function calls to complex Machine Learning algorithms such as Gradient Descent. Other libraries such as Keras will also be implemented for more complex modeling. [7]

3. METHODS

To begin the exploration into crypto currency market price prediction, we have implemented Linear Regression, a Supervised Machine Learning algorithm that outputs a linear prediction model. Linear Regression works by reducing the squared mean error of a proposed hypothesis □$\Theta(x)$ of what the actual output $y$is, which for our research is the market price of Bitcoin. The hypothesis function □$\Theta(x)$ is computed by multiplying the calculated matrix of "weights" called Θ, taking its transpose and performing matrix multiplication on x, which is a matrix of features such as the block chain size, Number of Bitcoin wallets, or Bitcoin Hash Rate. These features are vectors of known historic data that will be used to train our model to predict future prices given specific feature values

$$\square_{\theta}(x) = \Theta_0 + \Theta_1 x_1 . . . + \Theta_n x_n = \Theta^T x \quad \textbf{(1)}$$

To calculate the mean squared error of our hypothesis or more formally the cost, we take the difference of our hypothesis with the actual output y and square the value. This result is then summed over all the training examples m

of every feature n in x. We then normalize by dividing by two times the number of training examples m. In the case of one feature, the following cost function is shown (2).

$$J(\Theta) = 1/2m * \sum_{i=1}^{m}(h_{\Theta}(x_i) - y_i)^2 \quad \textbf{(2)}$$

Now that the cost function has successfully been computed, we have to iteratively reduce the mean squared error of our hypothesis by continually updating the matrix of weights Θ until they converge to global minimum value. This is performed by either using the Normal Equation, but for our purposes, we used Gradient Descent which is more efficient for larger data. This is achieved by taking the partial derivative with respect to Θ of our cost function, multiplying it by a learning rate α, whereby Θ will converge faster or slower depending on the chosen value of α, and subtracting this value from the current weight Θ. One can think of taking the partial derivative as finding what direction to move the current Θj value and increasing or decreasing the speed at which the process is performed by the learning rate α. The process is shown mathematically below

$$\boldsymbol{\Theta_j := \Theta_j - \alpha \frac{\delta}{\delta \Theta j}} \quad \textbf{(3)}$$

$$\boldsymbol{\Theta_j := \Theta_j - \alpha/m} * \sum_{i=1}^{m}(h_{\Theta}(x_i) - y_i)\boldsymbol{x_i} \quad \textbf{(4)}$$

The Next implemented strategy for crypto currencies like Bitcoin price prediction is using Recurrent Neural Networks (RNNs) with LSTM cells. Multi-feature RNNs with Long Short Term Memory or LSTM was developed to further improve prediction accuracy. Due to the RNN's back propagation feature, they are one of the best models which are often used for time-series prediction problems as they store the state of features from previous time-steps in memory and exploit them to predict the future. However, it has been proven time and again that RNN's performance drops when RNN have to remember information for a long period of time. LSTMs are special types of networks that solve the long-term dependency problem by means of a 'cell state'. The cell state allows for easy propagation of information across the entire chain.

4. EVALUATION

Finalizing the results of Linear Regression, Fig 1. Shows the result of performing Linear Regression using the day as a feature to predict the market price of Bitcoin over the last 365 days. Learning rate α was set to 0.03, and Gradient Descent has performed over 10,000 iterations. The cost function converged to a mean squared error of \$1039.58. Linear Regression has shown us that a simplified model of Bitcoin market price can be produced with minor complexity details, albeit larger error due to the constraint of the regression line being linear. As the price of Cryptocurrencies like Bitcoin exponentially begins to increase, the error of Linear Regression will continually get worse as a result of underfitting. Moving on to the second Linear Regression model, where Blockchain size is used as our main feature, we begin to find some striking similarities.

Blockchain size was used as a good representative feature for our model because increasing Blockchain size correlates to more investors making Bitcoin transactions, and Bitcoin market price is heavily weighed by consumer demand and interest. Shown in Fig 2 is the Linear Regression training output using Blockchain size as our feature with learning rate α at 0.3, and 10,000 iterations of Gradient Descent. After training our model, the cost function converged to a mean squared error of \$1053.29, almost the same as the previous date feature. This shows that date and Blockchain size have a dependency and strong correlation, therefore both features should not be used simultaneously when performing multi-feature modelling due to increased bias of theta weights when using features that vary in similar ways.

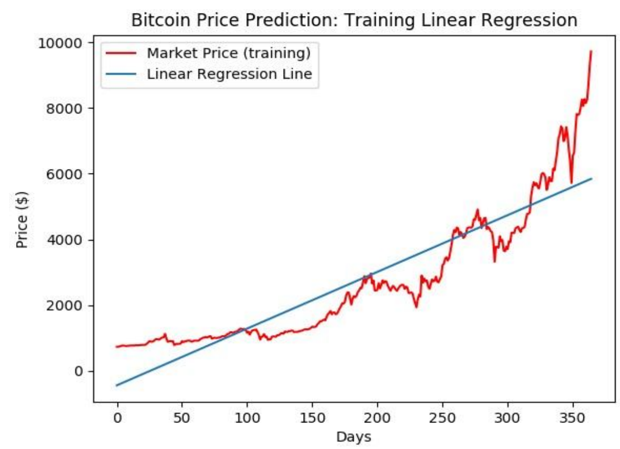

Figure 1: Linear Regression model given last 365 days, using training rate of 0.3 and 10,000 training iterations

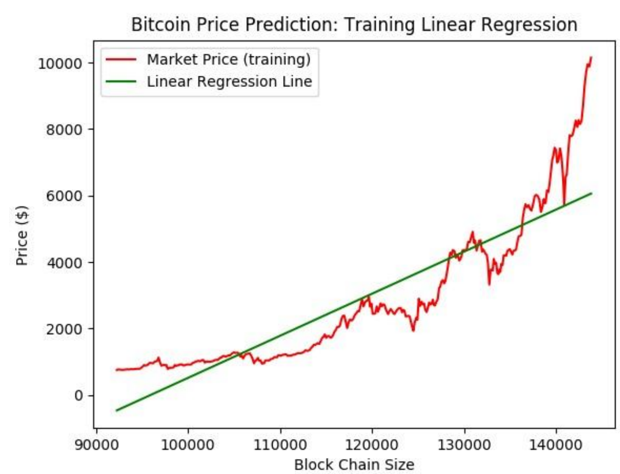

Figure 2:. Linear Regression model given Blockchain size, using training rate of 0.3 and 10,000 training iterations

Evaluating all Linear Regression models explored, Linear Regression with added features gave us the best results in terms of reducing the mean squared error of our training set, while also not overfitting to the data. The model was not affected by minute noise in the market price training set, which means it is generalizing to the data well, resulting in an improvement in accuracy given new unforeseen future feature data.

The Next model that we trained is the RNN with LSTM cells. The model was trained on the number of transactions per day, hashing rate, number of Bitcoin users per day, and transaction volume for every day over the past 3 years (except the last month, which was set aside for testing). Before the input features were supplied to the model, an additional feature selection optimization was applied. This was done to further reduce the feature set by removing input features from the model which had little correlation with the output. The feature selection method used was Reduced Feature Elimination (RFE). The idea behind RFE is straightforward. A simple model (the Linear Regression model in this case) is used as a training set with output values already known. Features are recursively removed and a model is built with the attributes remaining. Once the process is complete, features which do not significantly contribute to predicting the price of Bitcoin for the future can be removed.

After RFE was implemented (using a built-in function in the Scikit-learn library), the number of features was reduced from six to four. After feature selection was done, the LSTM model was tuned. To prevent over fitting, the number of neurons in the hidden layer had to be specified. Since there is no determining rule for picking the perfect number of neurons, the following popularly accepted rule of thumb was used:

$$^{Nsample}\alpha\ (Ninput+Noutput)$$

where α is a range of values from 2 to 10. For this case, α value of 3 was taken (the lower the value of α, the less the network overfits). Applying the above formula yielded a value close to 71

$$=\frac{1065}{3*(4+1)}=71$$

The number of epochs was chosen empirically to be 100. It can be seen in Figure 3 that by the time the epoch count becomes 100, both the training and test sets' loss values have plateaued and almost converged. This implies that the model was fitted. The model was trained on the data and tested against the actual prices during Fig 4 shows a plot of actual and predicted Bitcoin prices.

Given the percent accuracy results shown in Table 1, we see a percent accuracy of 69.9% when using the Linear Regression Model. [9] This can be attributed to the growing error when attempting to fit a linear model onto a very noisy, nonlinear data set. Moving to the Linear Regression model with features, we saw the better performance with a percent accuracy of 76.7%, which shows that increasing the features of the hypothesis function greatly reduces the cost function. [10] The Recurrent Neural Network model incorporating LSTM cells performed highly accurate than Linear Regression, having an accuracy of 96.2%. In regards to long term prediction making, we see RNN model begin to outperform the Linear Regression Model, which is more optimal for short term price predictions. [11] Referring to Figure 4, we see how the RNN model is sensitive to fluctuations of the bitcoin price and performs well in the short time frame. [8]

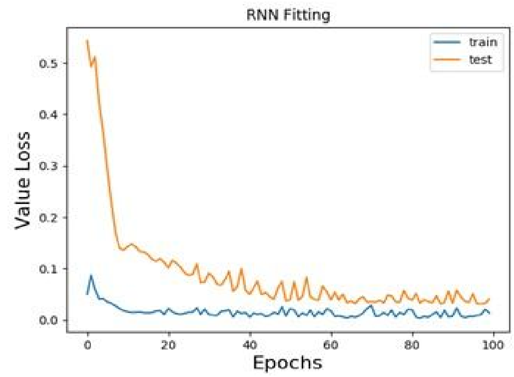

Figure 3: Value of the loss function for testing and training data sets vs the number of epochs

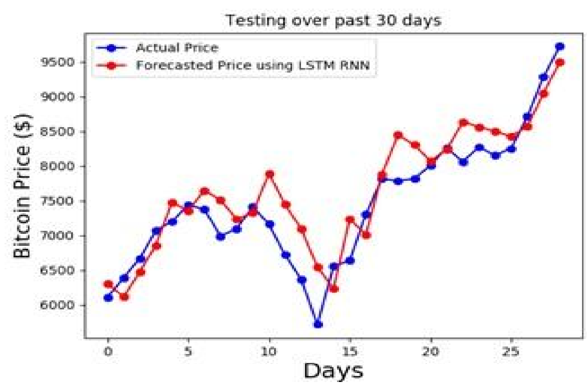

Figure 4: Actual vs predicted price of Bitcoin

The following table shows comparison of average prediction accuracies of the algorithms implemented in this project. Table 2 gives the used test data set for comparsion

Table 1: Bitcoin Price Prediction Accuracy over 15 Days

Machine Learning Model	Accuracy (%)
Linear Regression	69.9
Recurrent Neural Network With LSTM	76.7
Recurrent Neural Network With LSTM	96.2

Table 2: Test Dataset: Predicted vs. Actual Bitcoin Price

Linear Reg. ($)	Linear with Features Reg. ($)	RNN (LSTM) ($)	Market Price ($)
6036	7577	9510	9879
6013	7633	9581	9953
5992	7453	9792	9718
5970	7120	9168	9284
5949	6678	8686	8707
5929	6328	8637	8251

5. CONCLUSION

In this paper, several approaches for crypto currencies like Bitcoin price prediction were investigated. We compared the results of prediction with Linear Regression, Linear Regression with Features, and Recurrent Neural Networks with LSTM cells. The research contribution of this technique is that we predicted a numerical value of price instead of performing binary classification, as well as used multiple features to train the model. The LSTM method performed notably better than the other two approaches, and we believe that further research on using Neural Networks for time-series prediction is very promising to financial data analytics and other fields.

Our work can be extended further using some of the approaches described in the Related Work section. Namely, the LSTM-based model can be used as a part of the autonomous trading agent. It is worth investigating the scalability of our proposed approach. In particular, the important questions for further research are how far into the future should the price be predicted, and how many Bitcoins should be traded at a time by the autonomous agent.

The LSTM model, as well as the autonomous agent-based on it, can be further enhanced with sentiment analysis. Historical sentiments from Twitter, the number of search queries from Wikipedia and Google, and other metrics reflecting the public interest in Bitcoin can be used to influence the weights during model training. Moreover, the current sentiments can be combined with the prediction of the LSTM model to influence an autonomous trading agent's decision whether to buy or sell Bitcoins at a given moment of time.

Based on our research, it can be concluded that the future direction of work on predicting prices of stocks and crypto currencies has to take multiple metrics and features into account. Econometric data and sentiment analysis were only incorporated into projects that used simpler algorithms, such as Bayesian Regression. However, the projects involving DNNs or RNNs with LSTMs were trained on the datasets containing only the historical prices of stocks. This attests to the novelty of our work: we used multiple features in training the Neural Network, which is something not found in the research literature. Incorporating multiple features into prediction methods with these types of Neural Networks has the potential to produce more accurate results.

REFERENCES

1 "Bitcoin Price Index - Real-time Bitcoin Price Charts", CoinDesk Available: https://www.coindesk.com/price "
2. "A. Ng, "Linear Regression With Multi Variable", Stanford, CA"
3. "D. Nelson, A. Pereira and R. de Oliveira, "Stock Market's Price Movement Prediction With LSTM Neural Networks", in Neural Networks (IJCNN).
4. "M. Dixon, D. Klabjan and J. Bang, "Classification-based Financial Markets Prediction using Deep Neural Networks", Illinois Institute of Technology." https://doi.org/10.1162/neco.1997.9.8.1735
5. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780."
6. D. Shah and K. Zhang, "Bayesian regression and Bitcoin", *Massachusetts Institute of Technology*."
7. I.Georgoula, D. Pournarakis, C. Bilanakos, D. Sotiropoulos and G. Giaglis, "Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices".
8. Babitha, D., Ismail, M., Chowdhury, S., Govindaraj, R., & Prakash, K.B. (2020). Automated road safety surveillance system using hybrid cnn-lstm approach. International Journal of Advanced Trends in ComputerScience and Engineering, 9(2), 1767-1773. doi:10.30534/ijatcse/2020/132922020
9. Babitha, D., Jayasankar, T., Sriram, V. P., Sudhakar, S., &Prakash, K.B. (2020). Speech emotion recognition using state-of-art learning algorithms. International Journal of Advanced Trends in Computer Science and Engineering, 9(2), 1340-1345. doi:10.30534/ijatcse/2020/67922020
10. Bharadwaj, Y. S. S., Rajaram, P., Sriram, V. P., Sudhakar, S., Prakash, K. B. (2020). Effective handwritten digit recognition using deep convolution neural network. International Journal of Advanced

Trends in Computer Science and Engineering, 9(2), 1335-1339.doi:10.30534/ijatcse/2020/66922020
11. Enireddy, V., Gunda, K., Kalyani, N. L., & Prakash, K. B. (2020).Nature inspired binary grey wolf optimizer for feature selection in the detection of neurodegenerative (parkinson) disease. International Journal of Advanced Trends in Computer Science and Engineering, 9(3),3977-3987. doi:10.30534/ijatcse/2020/222932020
12. Ismail, M., Prakash, K. B., & Rao, M. N. (2018). Collaborative filtering-based recommendation of online social voting. International Journal of Engineering and Technology(UAE), 7(3), 1504-1507.doi:10.14419/ijet.v7i3.11630
13. Kolla, B. P., & Raman, A. R. (2019). Data engineered content extraction studies for indian web pages doi:10.1007/978-981-10-8055-5_45 Retrieved from www.scopus.com
14. Meghana, A. S., Sudhakar, S., Arumugam, G., Srinivasan, P., & Prakash,K. B. (2020). Age and gender prediction using convolution, resnet50 and inception resnetv2. International Journal of Advanced Trends in Computer Science and Engineering, 9(2) 1328-1334.doi:10.30534/ijatcse/2020/65922020
15. Prakash, K., Lakshmi Kalyani, N., Vadla, P. K., & Naga Pawan, Y. V. R.(2020). Analysis of mammography for identifying cancer cells using convolution neural networks. International Journal of Advanced Trends in Computer Science and Engineering, 9(2). 1184-1188.doi:10.30534/ijatcse/2020/44922020
16. Prakash, K. B. (2017). Content extraction studies using total distance algorithm. Paper presented at the Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology, iCATccT 2016, 673-679.doi:10.1109/ICATCCT.2016.7912085
17. Prakash, K. B. (2018). Information extraction in current Indian web documents. International Journal of Engineering and Technology (UAE),7(2.8), 68-71. Retrieved from www.scopus.com
18. Prakash, K. B. (2015). Mining issues in traditional Indian web documents. Indian Journal of Science and Technology, 8(32), 1-11 doi:10.17485/ijst/2015/v8i1/77056
19. Prakash, K. B., Dorai Rangaswamy, M. A., & Ananthan, T. V. (2014).Feature extraction studies in a heterogeneous web world. International Journal of Applied Engineering Research, 9(22), 16571-16579. Retrieved from www.scopus.com
20. Prakash, K. B., Dorai Rangaswamy, M. A., Ananthan, T. V., &Rajavarman, V. N. (2015). Information extraction in unstructured multilingual web documents. Indian Journal of Science and Technology,8(16) doi:10.17485/ijst/2015/v8i16/54252
21. Prakash, K. B., Eluri, R. K., Naidu, N. B., Nallamala, S. H., MishraP., Dharani, P. (2020). Accurate hand gesture recognition using cnn and rnn approaches. International Journal of Advanced Trends in Computer Science and Engineering, 9(3), 3216-3222.
22. Prakash, K. B., Kumar, K. S., & Rao, S. U. M. (2017). Content extraction issues in online web education. Paper presented at the Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology, iCATccT 2016, 680-685. doi:10.1109/ICATCCT.2016.7912086 Retrieved from www.scopus.com
23. Prakash, K. B., & Rajaraman, A. (2016). Mining of bilingual indian web documents. Paper presented at the Procedia Computer Science, 89 514-520. doi:10.1016/j.procs.2016.06.103
24. Prakash, K. B., Rajaraman, A., & Lakshmi, M. (2017). Complexities in developing multilingual on-line courses in the indian context. Paper presented at the Proceedings of the 2017 International Conference onBig Data Analytics and Computational Intelligence, ICBDACI 2017 339-342.doi:10.1109/ICBDACI.2017.8070860
25. Prakash, K. B., Rajaraman, A., Perumal, T., &Kolla, P. (2016).Foundations to frontiers of big data analytics. Paper presented at the Proceedings of the 2016 2nd International Conference on Contemporary Computing and Informatics, IC3I 2016, 242-247. doi:10.1109/IC3I.2016.7917968
26. Prakash, K. B., Rangaswamy, M. A. D., & Raja Raman, A. (2012). ANN for multi-lingual regional web communication doi:10.1007/978-3-642-34500-5_56
27. Prakash, K. B., Rangaswamy, M. A. D., & Raman, A. R. (2012).Statistical interpretation for mining hybrid regional web Documents doi:10.1007/978-3-642-31686-9_58
28. Reddy, A. V., Vege, H. K., & Prakash, K. B. (2020). Efficient and accurate hybrid deep learning model for multimodal disease risk prediction. International Journal of Advanced Trends in Computer Science and Engineering, 9(2), 1262-1267. doi:10.30534/ijatcse/2020/55922020
29. Vadla, P. K., & Prakash, K. B. (2020). Residue based adaptive resource provisioning through multi-criteria decision and horizontal scaling of vm's in agent-based model for federated cloud. International Journal of Advanced Trends in Computer Science and Engineering, 9(2). 1610-1622.doi:10.30534/ijatcse/2020/108922020