J. Electrical Systems 20-4s (2024): 539-553

[1] **Amal Saad Alshehri***

Predicting Cryptocurrency Returns Using Classification and Regression Machine Learning Model

Abstract: - People are starting to see the cryptocurrency market as a viable source of income and investment, similar to the stock market, as the concept of cryptocurrencies continues to gain popularity. Predicting Bitcoin returns is related to financial machine learning, which uses time series to forecast price variance. This study starts with the daily close price of Bitcoin for its initial dataset. The price is transformed into percentages and binary classes, which categorize into "Up" and "Down", after which a time series is applied to produce two datasets: a categorical dataset for classification and a numerical dataset for regression. For classification that represents a Binary classification in asset-price forecasting, k-fold cross-validation is applied to ensure that the best classifiers are selected for testing and analysis. Most of the regression analysis was based on visualization, which displayed the predicted prices by each regressor in front of the original values and helped analyse the models' results more accurately. The outcomes of this study were achieved by anticipating bitcoin returns using classification and regression machine learning models, despite the approaches' low accuracy and significant precision rate to the "Up" class. At this stage, with a significant limitation regarding the dataset and a lack of other indicators, a model capable of predicting future variations is considered a beneficial addition for many trading tools or even for crypto market analysts.

Keywords: Bitcoin predictability, Time-series cross validation, Binary classification in asset-price forecasting, Financial machine learning.

I. INTRODUCTION

As a significant player in the global financial landscape, cryptocurrencies have gained the interest of regulators, governmental organisations, institutional and individual investors, researchers, and the general public. Cryptocurrencies are brand-new money that is sweeping the financial industry and catching the attention of industry pioneers. Investors and industry professionals are both concerned about making accurate predictions about the values of the cryptocurrency market because prices are rising quickly and responding similarly to stock market price changes.

Although this market exhibits similar behaviours to other stock markets in terms of expected volatility, investor confidence has been reflected in it [1,2,3]. Even though prices are quite volatile, and it is hard to identify certain crucial factors and quantify the extent to which they affect the price, researchers' efforts in the area of currency price forecasting continue in an effort to understand the digital financial landscape. There is still no conclusive theory explaining how to price Cryptocurrencies [4]. A study has shown that cryptocurrency price is positively and statistically significantly correlated with computing power and network adoption; the study may also be a tool for identifying additional factors such as regulatory and political risks affecting the returns on these digital assets [5]. The volatility of cryptocurrency exchange values is mostly brought on by market emotions, a lack of government regulatory oversight, and the fact that these currencies lack an intrinsic value. This cryptocurrency's popularity is one of the key factors influencing its price. One of the critical causes of price instability and volatility is the ineffective integration of cryptocurrencies into the traditional currency and macroeconomic markets [4].

The price of cryptocurrencies does not significantly depend on societal factors [6]. More research should be done in this area to get better findings with more reliability. With the exception of bitcoin, all cryptocurrencies are referred to as "altcoins" which is a term that originates from the notion that these coins are alternatives to bitcoin. There is a plethora of other altcoins available, but before any of them were made, Bitcoin was the first and most valuable asset. These alternative currencies aim to improve or add advantages like quick transactions and low energy consumption. The cryptocurrency sector is nearly new and still developing, Bitcoin holds the distinction of being the first cryptocurrency or asset. One of the most critical leading indications of the performance of cryptocurrencies is the performance of Bitcoin. The dominance component is essential because Bitcoin prices influence markets

through their dominance. The market capitalization of Bitcoin rules the cryptocurrency industry and significantly influences it. Altcoins do not have the same level of widespread acceptance as Bitcoin. They also

[1] Jazan University
* Corresponding Author Email: ashahri@jazanu.edu.sa

carry most of Bitcoin's traits because it was developed using Bitcoin's source code as a starting point and then improved upon [7]. Prices may differ somewhat between Bitcoin and its clones, but they often do not. Numerous scholars are interested in investigating the connection and impact between Bitcoin and other altcoin markets. After analysing the performance of the daily data of 16 alternative currencies in addition to Bitcoin between 2013 and 2016, this study [8] concluded that there is, in fact, a significant correlation between the prices of Bitcoin and other alternative currencies, with the correlation being more evident in the short term. The study found that, compared to bitcoin prices, macro-finance factors have a longer-term effect on altcoin pricing. In contrast to stock markets, cryptocurrencies have a strong connection.

It is not advised to create a portfolio using Bitcoin and other alternative currencies since the value of these currencies is significantly impacted by the price of Bitcoin [9]. The price of other encrypted digital currencies is significantly influenced by Bitcoin, the first and most important symbol for the whole cryptocurrency industry. Various factors contribute to this, including that the most well-known alternative currencies are just enhanced or expanded versions of Bitcoin. The liquidity of Bitcoin also contributes to its continued dominance. Still regarded as the safest cryptocurrency asset is bitcoin. It is impossible to accept or refute the claim that Bitcoin has a 100% impact on the rates of other alternative currencies. In other words, because there are many alternative cryptocurrencies and the financial sector is intricate, the rise in bitcoin prices does not always grow by the same percentage as all alternative cryptocurrency rates. The expected drop or rise may not occur because certain elements seem to cancel out the effects of other ones. Bitcoin and other cryptocurrencies are, nonetheless, closely connected.

II. Research Aim

The aim of this research is to investigate the feasibility of utilizing machine learning models for predicting specific cryptocurrency returns using classification and regression models.

III. Research Objectives

1- To collect data from Yahoo Finance API and build datasets for classification and regression models.
2- To evaluate different classifiers, including SVM, XGBoost, and RandomForest, and identify the best-performing model.
3- To compare the results of the classification and regression models to determine which method is better at predicting cryptocurrency price movements.
4- To investigate the use of ensemble machine learning methods for predicting cryptocurrency market returns.

IV. Research Contribution

This research contributes to the field of finance and machine learning by providing insight into the feasibility of utilizing machine learning models for predicting cryptocurrency returns. The research identifies the best-performing classification and regression models and highlights the importance of feature selection when building datasets. The investigation of ensemble machine learning methods for predicting cryptocurrency market returns also adds to the existing literature on the topic.

V. Paper Organization

The paper follows a systematic structure, starting with the research methodology. It then outlines the tools for data collection and analysis, presents the results, and discusses ethical considerations. Limitations of the study are outlined, followed by a concise analysis of findings. The paper concludes by suggesting future research directions.

VI. Methodology

The methodology approach followed by this project is Machine Learning Conceptual Modelling. It is a modern approach introduced in many papers and consists of pairing conceptual modelling with machine learning because both have long been recognized as essential research areas [10,11]. The conceptual model consists of a well-determined development process that starts with problem-understanding and ends with Analytical decision-making. Taking into account the fact that machine learning problems might vary depending on the case study included in the project, starting with the overall development process, a tailored method was created to meet the needs of this project, as shown in fig. 1 The first parts consist of data-related parts, which start with data collection from a reliable data source such as yahoo finance API and then carrying out the required pre-processing, cleaning, and feature engineering of the collected data. Time-series analysis will be performed on the data, which will be our features for the following parts. Instead of relying just on one discipline, such as regression or classification, both will be applied in accordance with the dataset, and two more sub-datasets will be created. The first will be a continuous values dataset labelled with the daily per cent change in price for the chosen cryptocurrency (daily closing price was converted to daily price percent change). The second dataset will also include categorical data, which will be divided into the classifications “Up” and “Down” based on the daily price percent change, which represents price variation.

For the classification problem, cross-validation will be used to compare the most popular classification method, and for the regression problem, a collection of linear and non-linear regression algorithms will be used. The deep learning model will be approached differently since the layers’ development, and the architecture as a whole will be the core issues of concern.

Multiple training will be performed then, followed by the testing and validation part. The Result analysis comes just before validating the final model and setting it for deployment, so it includes improved visualization compared to the actual data (only for regression) and then comparing both predictions from the various trained classification and regression models.

Regression and classification employ several evaluation methods and metrics.

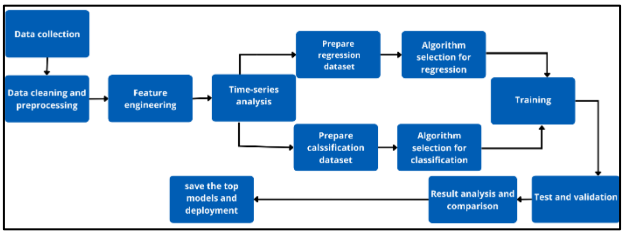

Fig. 1. Machine learning development process

VII. Tools and Resources

Several tools and resources were employed for this research, starting with Python as the main programming language because it is the most common language for data science and offers a large number of packages that can be used for many different computer science and data science fields. In terms of data collection, Bitcoin-related data was gathered via the Yahoo Finance API. Since it is a reliable source of data with many benefits the most well-known of which is that it has already been implemented as a Python package, by just installing the package and simple lines of code, the data will be ready for use in subsequent phases without the need for an API-Key or with limited access. Tensorflow, Keras, Sickit-Learn, Pandas, Numpy, Matplotlib, Plotly, and Dash, are the primary tools utilised; these are the essential packages for our machine learning conceptual model. The working environment used was Google Colab, a cloud-based alternative to Jupyter Notebook that removes most hardware issues and other potential limitations while working on data science research. To guarantee the optimum development circumstances for the project, every resource and tool chosen for it was carefully considered after extensive testing with various tools. The collected data will be for Bitcoin since Bitcoin influences the majority of the volatility in the cryptocurrency market. The initial dataset will contain these features: Open, High, Low, Close, Adj Close, and Volume. The dataset date range will be from “2014-09-17” (which is the minimum date you can start within the API) to “2022-07-30” (a randomly picked data to have a specified end date), which comprises 2874 days (row of data). The initial dataset is shown in fig. 2 as follows:

Date	Open	High	Low	Close	Adj Close	Volume
2014-09-17	465.864014	468.174011	452.421997	457.334015	457.334015	21056800
2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	34483200
2014-09-19	424.102997	427.834991	384.532013	394.795990	394.795990	37919700
2014-09-20	394.673004	423.295990	389.882996	408.903992	408.903992	36863600
2014-09-21	408.084991	412.425995	393.181000	398.821014	398.821014	26580100
2014-09-22	399.100006	406.915985	397.130005	402.152008	402.152008	24127600
2014-09-23	402.092010	441.557007	396.196991	435.790985	435.790985	45099500
2014-09-24	435.751007	436.112000	421.131989	423.204987	423.204987	30627700
2014-09-25	423.156006	423.519989	409.467987	411.574005	411.574005	26814400
2014-09-26	411.428986	414.937988	400.009003	404.424988	404.424988	21460800

Fig. 2. Machine learning development process

In light of the fact that the forecast would be based on daily variation and that the closing price is the best match to be used as the daily price, all the columns were then dropped, with the exception of the close price column. Since the project idea does not involve non-stationary values like prices, a new column named "Close%" was created. This column will be utilised for the regression portion of the project and reflects the daily percent changes for Bitcoin. A second column called “Variation" was also produced using the column “Close%” with values 1 and 0. The “Variation" column denotes the price percent indications (1 for “Up” and 0 for “Down”). The modified dataset is also displayed in fig. 3 below:

Date	variation	Close %
2014-09-17	0	0.000000
2014-09-18	0	-7.192558
2014-09-19	0	-6.984265
2014-09-20	1	3.573492
2014-09-21	0	-2.465854
2014-09-22	1	0.835210
2014-09-23	1	8.364742
2014-09-24	0	-2.888081
2014-09-25	0	-2.748309
2014-09-26	0	-1.736994

Fig. 3. Updated Dataset

For more illustration, fig. 4 shows how the variation column distribution looks like between the two classes “Up” and “Down”:

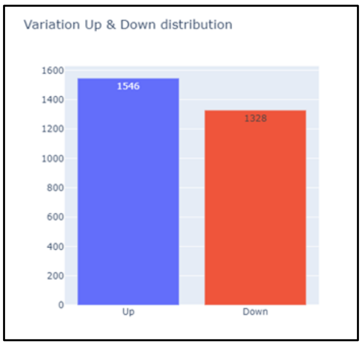

Fig. 4. Variation column classes distribution Variation column classes distribution

We implemented the following in the data preprocessing section:

1- Changing the date column to use the pandas DateTime format.
2- Normalizing the price percent change (the “Close%” column).
3- Feature generation utilizing time-series analysis using a custom function on the provided dataset based on a 21-time step, which is accomplished by extracting the features from the original time series at each time step while considering a predefined number of past values.

A rolling mechanism will provide a sub-time series of the last (m) time steps to build the features. The final part will consist of splitting into two sub-datasets, one for regression and the other for classification, which mainly have the same features but different labels, then split these datasets into 70% train and 30% test. Regression and classification are the two main sorts of issues used in this research. Although each has its own formulation, they both fall under the category of supervised learning, meaning that the prediction will be based on a labelled dataset. In this case, classification and regression both use the same input feature data, but they differ when it comes to the output/label data, which must be continuous for regression and categorical or binary for classification. The input data will be defined by a set of time-series vectors extracted after the feature engineering part. Each data point consists of 21-time steps, which present the feature of that specific data point built based on the last steps, basically a translation for the daily price percent change variation. For classification, the output vector will be made up of two primary classes: 1, which represents the “Up” class and denotes that the daily price percent change will go up, and 0, which denotes the “Down” class and indicates that the daily price percent change will go down. The output vector for regression will include the daily price percent change for each day, which ranges from -1 to 1. (normalized). For each part a group of models will be trained to predict the possible classes or values for each data point, after which the model will be able to make future predictions on new cases.

VIII. RESULTS AND ANALYSIS

8.1 Prediction Results for classification:

This section discusses the results of the classification prediction models in addition to the regression models. The primary label for classification was either the price percent was positive or negative for that day, representing a data point. For regression, the label will be the daily price percent change. 70% of the data were utilised for training, and 30% were used for testing. However, in the case of the ANN, 40% of the testing set was used for validation while training. As each approach has its own criteria for model evaluation, the results will first be analysed for the classification component, followed by the regression component.

8.1.1 *Classifiers comparison:*

The confusion matrix allows for representing the performance of a classification model on the test set that will help determine the other metrics like accuracy, precision and recall. All the score calculation was done using the 30% test set on the models. Cross-validation, a resampling technique that tests and trains a model at different iterations using different parts of the data, is the first step in analysing the results for machine learning models. It runs on the entire dataset and is based on the number of folds (8 folds) that make up the number of years in the dataset. It is primarily used in cases where the goal is prediction and can measure how accurately a predictive model can perform on different portions of the dataset. A collection of machine learning classifiers will be utilised for the comparison of the result, including: Logistic Regression, Linear Discriminant Analysis, KNN, Decision Tree, GaussianNB, Linear SVM andSVM, Random Forest, LGBM and XGBoost. Figure \ref{fig:BoxPlots}, which shows the range of accuracy provided by each classifier after cross-validating each technique for eight folds run, is shown below:

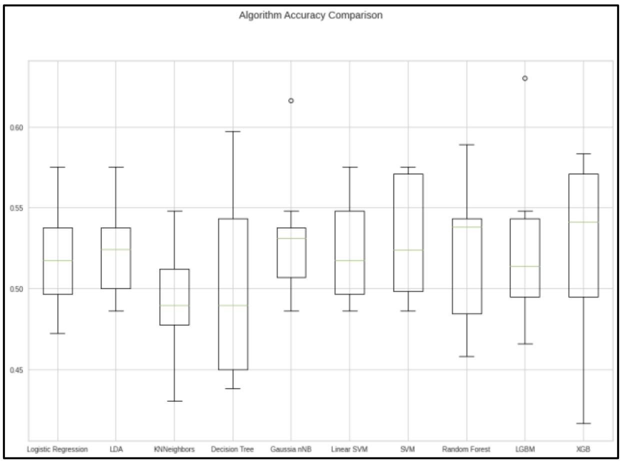

Fig. 5. Box Plots for classifiers accuracy range comparison

To determine which classifiers can achieve high accuracy throughout this procedure, just the maximum accuracy for the eight folds training/testing will be shown, as shown in fig. 6 below:

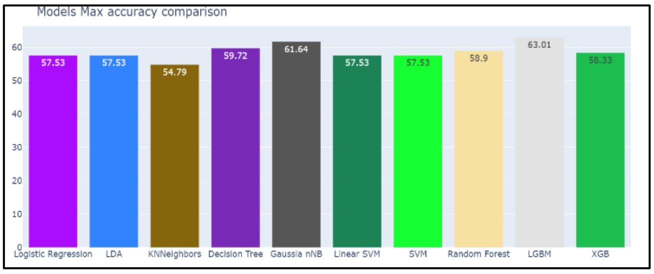

Fig. 6. Bar Plot for classifiers max accuracy comparison

Bellow, the average accuracy results for each classifier are shown. Based on all the results, it is possible that the choice will not only be based on accuracy scores but also on consistency and other factors. In order to get the highest accuracy, the selected classifiers will also undergo hyper-parameter tuning; the resulting average accuracies are as follows:

-Logistic Regression: 51.88%.

-LDA: 52.40%

-KNNeighbors: 49.47%

-Decision Tree: 50.36%

-Gaussia nNB: 53.08%

-Linear SVM: 52.23%

-SVM: 53.09%

-Random Forest: 52.23%

-LGBM: 52.41%

-XGB: 52.58%

Based on all the findings from this step, this collection of classifiers will be chosen for the following reasons:

- -SVM: Because it has the highest average accuracy.

- -XGBoost: The most recommended approach for machine learning classification models, even if earlier results were not particularly strong.
- -RandomForest: This model allows for the highest accuracy obtained through hyper-parameter adjustment.
- -A custom ANN model will be created and included in the analysis and comparison in addition to these classifiers.

8.1.2 Classification Experiments Results:

After executing the training for each of the selected models, all the metrics will be computed in this part. Then analysed and evaluated. The confusion matrix, classification report, precision, recall curve, and roc curve will be illustrated. After that, all the models will be compared globally.

SVM result analysis:

SVM evaluation metrics gave these values:

-Accuracy: 48.99%

-Balanced Accuracy: 47.68%

-Precision: 65.62%

-Recall: 51.41%

-F1: 57.65%

The confusion matrix for the SVM model is shown in fig. 7 below:

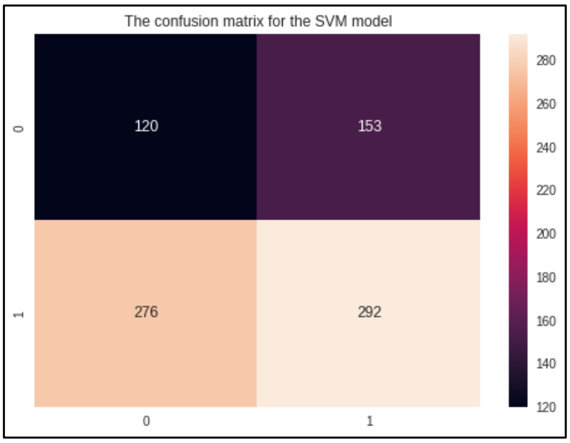

Fig. 7. SVM confusion matrix

The classification report, which is used to evaluate the accuracy of predictions made using a classification model, is illustrated in fig. 8. It includes all the necessary metrics for each class, such as accuracy, precision, recall, and the f1-score.

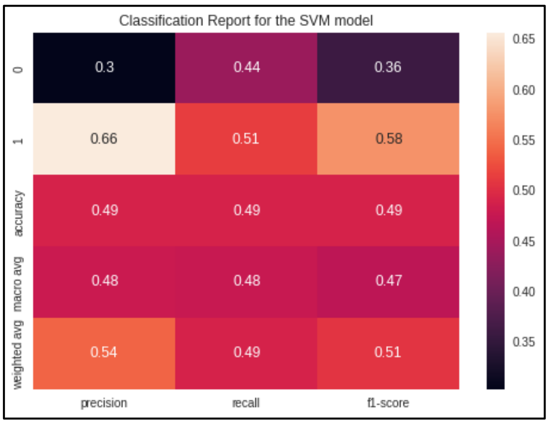

Fig. 8. SVM classification report

The precision-recall curve plots the trade-off between precision and recalls for different thresholds, and this is the curve in the case of the SVM model shown in fig. 9 below:

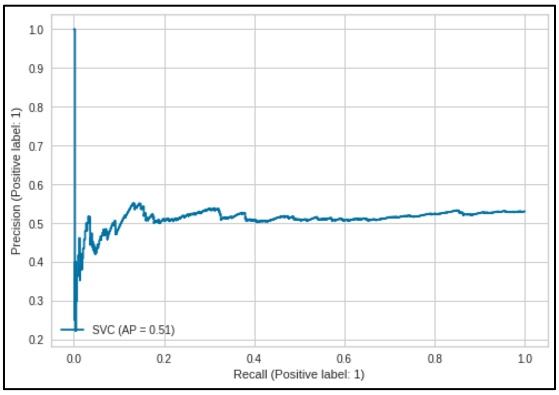

Fig. 9. SVM precision-recall curve

The ROC curve (Receiver Operating Characteristic Curve) is a plot showing the performance of a classification model at all classification thresholds, as shown in fig. 10 below in the case of the SVM model.

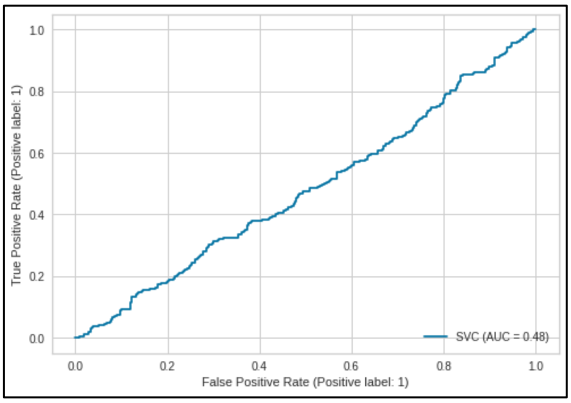

Fig. 10. SVM ROC curve

In conclusion, the SVM model does not provide the expected results as it is lower than the minimum required accuracy for this kind of classifier which is 50%.

XGBoost result analysis:

The XGBoost classifier evaluation metrics results:

-Accuracy: 52.56%

-Balanced Accuracy: 51.68%

-Precision: 70.70%

-Recall: 53.94%

-F1: 61.22%

A detailed analysis of results in the form of confusion metrics, precision-recall, and ROC curve is provided in fig. 11,12,13,14.

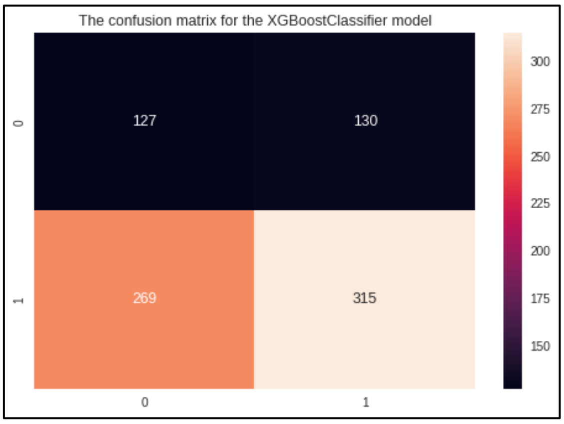

Fig. 11. XGBoost confusion matrix

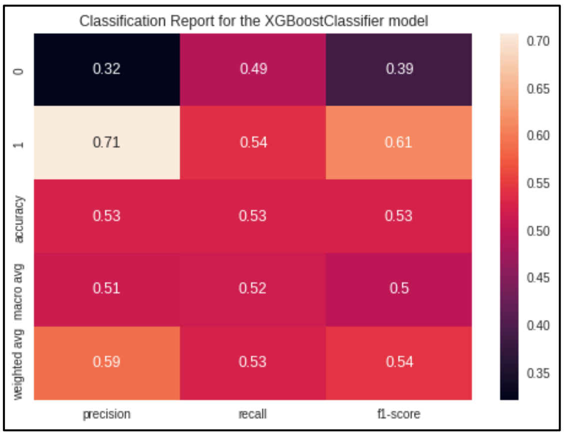

Fig. 12. XGBoost Classification report

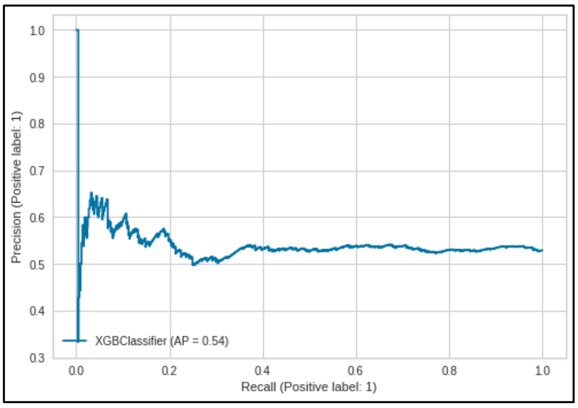

Fig. 13. XGBoost Precision-Recall curve

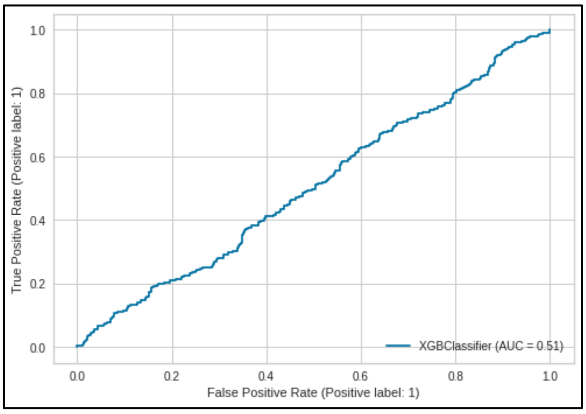

Fig. 14. XGBoost ROC curve

The XGBoost classifier seems to have the highest accuracy and a value considerably acceptable, especially to be implemented into a trading strategy or a trading bot.

Random Forest Result Analysis:

For Random Forest, a typical run was performed at the start, but then hyper-parameters tuning was applied to get the best possible parameters from the model, which will get the best accuracy score possible, starting with the evaluation metrics results for the first run:

-Accuracy: 53.51%

-Balanced Accuracy: 52.87%

-Precision: 86.07%

-Recall: 53.79%

-F1: 66.21%

A detailed analysis of results in the form of confusion metrics, precision-recall, and ROC curve is provided in fig. 15,16,17,18.

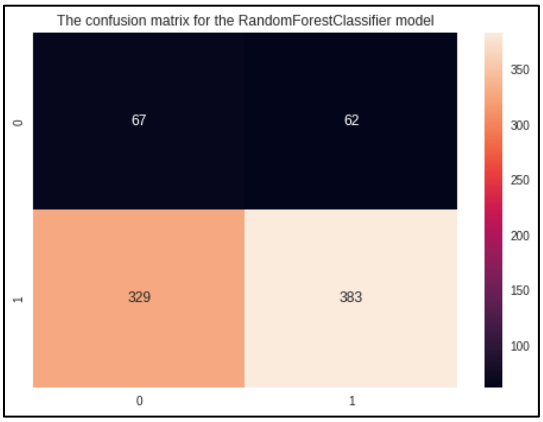

Fig. 15. Random Forest confusion matrix

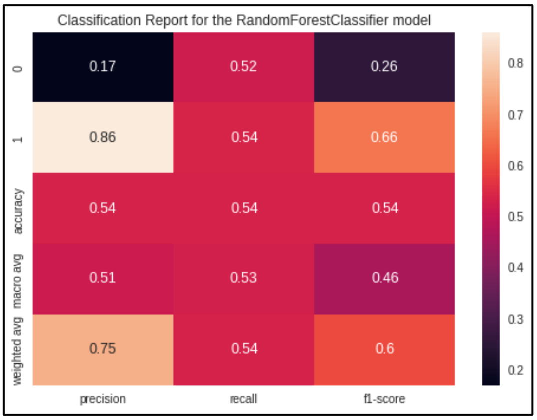

Fig. 16. Random Forest classification report

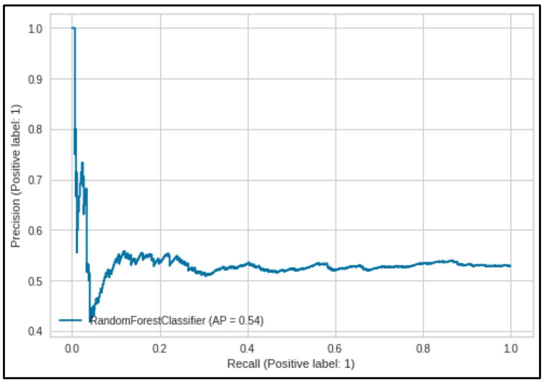

Fig. 17. Random Forest precision-recall curve

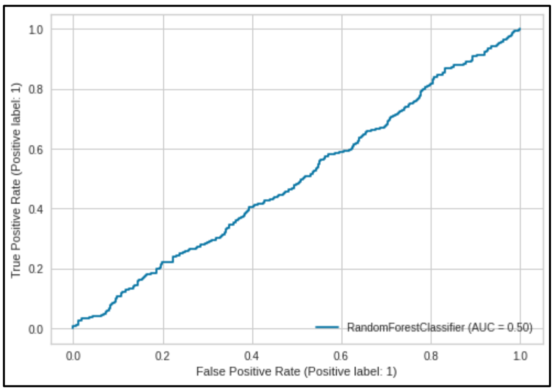

Fig. 18. Random Forest ROC curve

In addition to these results, thanks to hyperparameters tuning, the best estimator was able to hit 54.65% accuracy, as it seems that the Random Forest classifier performed the best among all the machine learning classifiers chosen. However, the only drawback is that this classifier has a very high precision rate for class "1" compared to other models with lower accuracy, which have a closer rate between the two classes.

ANN result analysis:

The result analysis for the ANN model has some similar points to the other machine learning models. However, ANN will have other comparison aspects, as the focus will be on the loss and accuracy. In addition to the validation set, which validates each training epoch and calculates the validation loss and validation accuracy to help to notice any over/underfitting during the training. Also, an extra function was added that helps save the best model during the whole training, which is ModelCheckpoint, to save the model with the lowest "validation loss". EarlyStopping was added and tested. However, it seemed that it is not helpful for this case, so the focus will be only on the ModelCheckpoint method. Fig. 19 shows the classification report. The results of evaluation metrics for ANN:

-Accuracy: 52.38%

-Balanced Accuracy: 50.23%

-Precision: 85.82%

-Recall: 53.24%

-F1: 65.71%

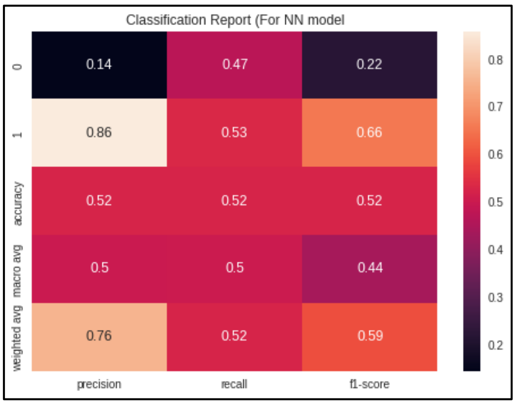

Fig. 19. ANN Classification Report

Fig. 20, 21 will plot the training and validation loss/accuracy as a function of the epoch during the model training, which helps analyse the training and look for any over/underfitting. In this case, the train loss/accuracy will be in the purple line, and the validation loss/accuracy will be in the orange line.

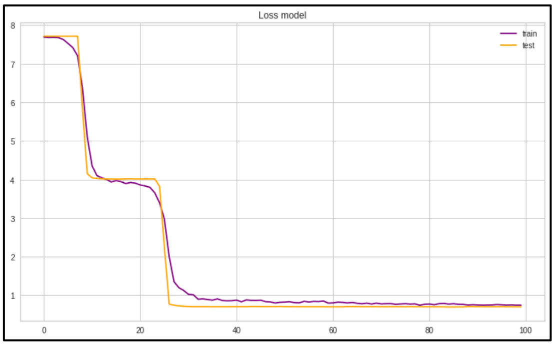

Fig. 20. ANN Loss plot in the function of epoch

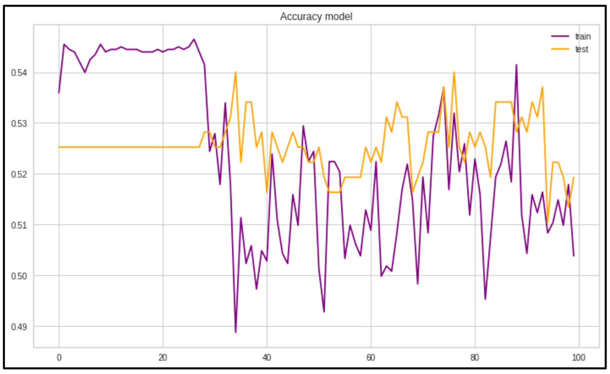

Fig. 21. ANN Accuracy plot in the function of epoch

It can be observed from the figures above that the result does not get better even when using a good build ANN model.

In the table 1 we demonstrate **the final classifiers global comparison** (green colour means the highest value in the column, yellow colour means a considered model)

Table 1. Evaluation metrics comparison for classification models

Models	Accuracy	Balanced acc	Precision	Recall	F1-score	AUC-ROC
SVM	48.99%	47.68%	65.62%	51.41%	57.65%	48.00%
XGBoost	52.56%	51.68%	70.79%	53.94%	61.22%	51.00%
RandomForest	53.51%	52.87%	86.07%	53.79%	66.21%	50.00%
ANN	52.38%	50.23%	85.82%	53.24%	65.71%	NA

The major noticed point from prior confusion matrices and classification reports is that most classifiers have high precision rates for class (1), which results in high precisions and low accuracies, reducing the model's efficiency. This prompts us to consider not only the high accurate model but the more stable and balanced model, which are the XGBoost model and the Random Forest model.

8.2 Prediction Results for regression:

The linear model, ensemble regression models, and RNN models make up the three primary components of the regression section. The same time-series analysis feature utilised for classification will be the input for all models, and the intended result will be the daily price percent change. In order to determine how accurate the model is at forecasting the daily percent change, the assessment will be based on three metrics (MAE, MSE, and RMSE) in addition to a comparison plot that displays the results between the projected percentages and the actual percentage.

Linear Regression Model:

Results of Model evaluation metrics for Linear Regression (RMSE, MSE, MAE):

-Mean Absolute Error - MAE: 8.58%

-Mean squared Error - MSE: 1.44%

-Root Mean squared Error – RMSE: 12.02%

The prediction on the training set is shown in the red line, the prediction on the testing set is shown in the green line, and the actual daily price percent changes are shown in the blue line in the following fig. 22. It appears that the linear regression model was unable to produce a value that was close to what the actual results should be.

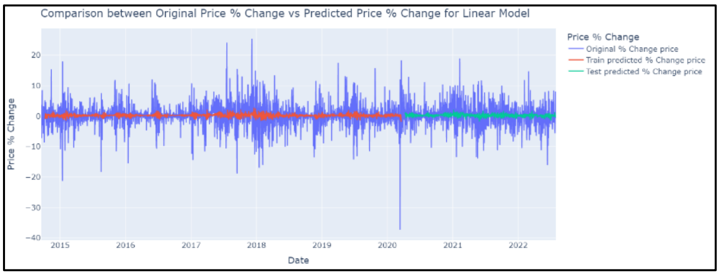

Fig. 22. Comparison between original and Predicted values for Linear Regression

Ensemble Regression Model(s):

XGB Regressor Model:

Model Evaluation metrices RMSE, MSE, MAE (for XGB Regressor):

-Mean Absolute Error - MAE: 9.65%

-Mean squared Error - MSE: 1.67%

-Root Mean squared Error – RMSE: 12.92%

As plotted in the fig.23 below, XGBoost Regressor could have a very close prediction on the training set and a great result regarding the testing set.

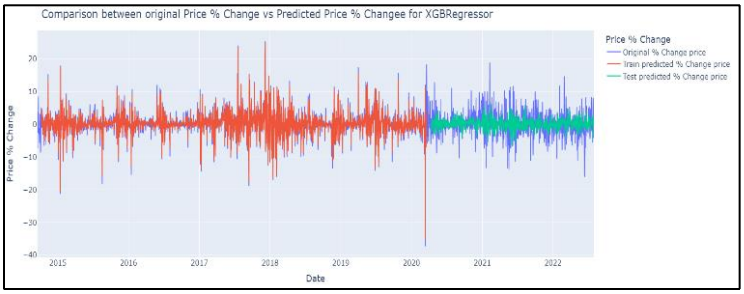

Fig. 23. Comparison between original and predicted values for XGB Regressor

Random Forest Regressor Model:

Model Evaluation metrices RMSE, MSE, MAE (for Random Forest Regressor):

-Mean Absolute Error - MAE: 8.66%

-Mean squared Error - MSE: 1.44%

-Root Mean squared Error – RMSE: 12.00%

The results from the Random Forest model are not that great compared to the previous ensemble model, as shown in fig. 24 below:

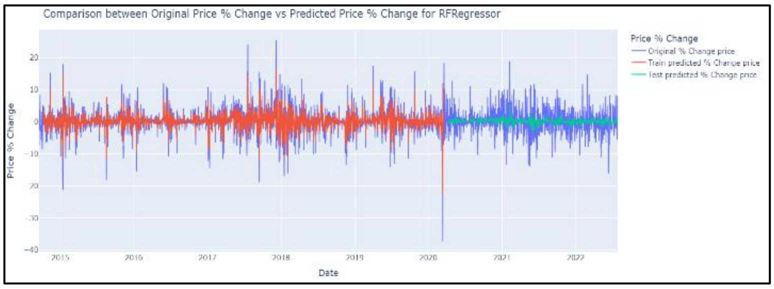

Fig. 24. Comparison between original and predicted values for Random Forest Regressor

RNN Regression Model (LSTM):

Model Evaluation metrics RMSE, MSE, MAE results (for LSTM):

-Mean Absolute Error - MAE: 3.70%

-Mean squared Error - MSE: 13.69%

-Mean squared Error - RMSE: 2.64%

Despite having a lower loss value, the LSTM model provides very far results, as shown in the fig. 25:

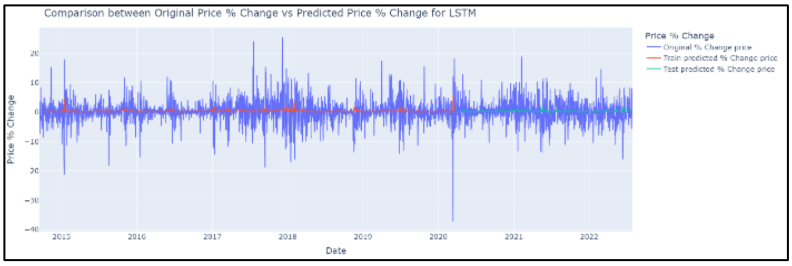

Fig. 25. Comparison between original and predicted values for LSTM

By attempting to predict the price returns using traditional regression models and various classification approaches, all experiments have been done to test all the potential models that might be employed in such a field. The Random Forest Classifier and the XGBoost Classifier are the best models for classification. Then, for regression, it demonstrates that the XGBoost regressor could be able to predict the anticipated values rather well. The major thing that stands out is that XGBoost is the best model in both classification and regression because it is known for its ability to get the best performance for boosted tree algorithms and great computational speed.

8.3 Comparing and merging Classification and Regression models:

The only way to compare regression models to classification models is to transform the numerical prediction results of the regression into categorical results that can be divided into two classes, “1” or “0” By taking the prediction values and changing them into “1” if the value is positive or “0” if the value is negative, it will be possible to calculate accuracy using those transformed results. In addition, the accuracy will be calculated using the testing set first and then using the entire dataset and these are the results for the testing set:

-RandomForest Classifier: 53.51%
-XGBoost Classifier: 52.56%
-RandomForest Regressor: 52.91%
-XGBoost Regressor: 52.68%
For the full dataset results:
-RandomForest Classifier: 66.83%
-XGBoost Classifier: 65.15%
-RandomForest Regressor: 54.00%
-XGBoost Regressor: 35.93%

Moving on to another part, we combine all the models into one predictor that utilises a majority voting system, gathering the predictor results into one array and beginning to iterate data point per data point. By using the majority voting system, the decision will be made based on the class that receives the most votes, and by doing this, we will be able to predict the outcomes of five predictions that came from various models which can be used to calculate the accuracy value and these the values we got:

-Calculating with the testing set only: 53.27%
-Calculating with the full dataset: 65.43%

And the following fig. 26 shows how the voting process actually works:

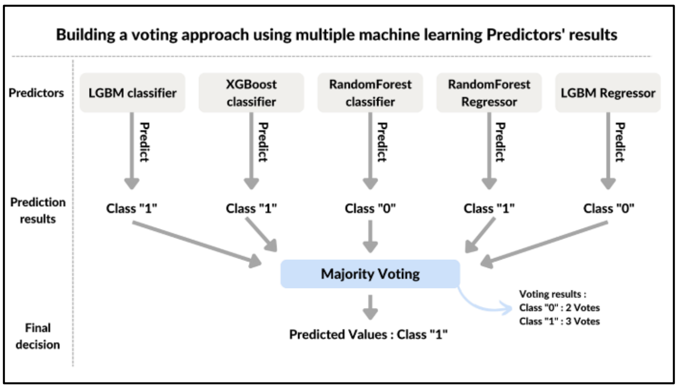

Fig. 26. Explaining the created voting system that merges multiple classification and regression models into one predictor

In conclusion, classification models offered higher accuracy than regression models. The combination method used can be improved in future work by adding more prediction models or improving the voting system and can be easily transformed into a reliable predictor more efficient than just one machine learning model.

IX. ETHICAL, SOCIAL, AND LEGAL ISSUES:

Cryptocurrencies are a widespread global economic phenomenon, which has repercussions at the economic and individual levels, as it affects the monetary policies of the state in terms of the money supply due to the lack of state control over its issuance, and also affects the financial policy to facilitate the process of tax evasion, as well as the payments and credit system due to the absence of mediation. However, at the same time, it was distinguished by its wide acceptance due to the dealers' confidence due to its fluidity in dealing and benefiting from it in all transactions efficiently and quickly through the widespread trading platforms. On the other hand, there are significant risks to investing in Bitcoin and cryptocurrencies in general, as volatility is a crucial characteristic of cryptocurrencies. The price of Bitcoin and cryptocurrencies, in general, is incredibly volatile because it is a very young and emerging market. It is common for the price of Bitcoin to experience sharp fluctuations within a day or even minutes. Therefore, *this project is only intended for research purposes, and we cannot recommend using its results to make an investment or speculative decisions in cryptocurrencies*.

X. LIMITATIONS AND PROBLEMS:

One of the significant limitations of the project was the dataset used for the training because of the lack of external features, which makes most of the work based on the time-series features, which were analysis for the daily price percent variation for Bitcoin for 2874 days, this is a significant factor for having a low accuracy value and most of the classifiers get distracted by predicting a single class. However, the overall approaches provided reliable models that can be integrated into a useful trading tool. In the same matter, a lot of other factors can be related to cryptocurrency price variation; some of them is unpredictable, like economic crisis or wars, and other different factors that can be suitable for the prediction that was not experimented with in this project and because most of them are beyond the scope of this project. However, things are a bit different regarding the cryptocurrency market for various reasons, one of which is that blockchain and cryptocurrencies are still mostly unheard of, especially when talking about the Web 3.0 era, the future of the web. On the other hand, the comprehension of cryptocurrencies and forecasting models may be advanced by extensive research and analysis, which will improve as more data, features, and other factors become accessible.

XI. Evaluation and Discussion:

According to all of the analysis and metrics calculations, whether for the classification or regression approach, the overall result is that the average accuracy for both is about 53\% (precisely 52.91\%) based on over 60\% precision rate for the class “1" (calculated only for classification models), which represents that the price would increase in that day. The RandomForest classifier was the top-performing model, outperforming both classification and regression models with an accuracy rate of 53.51\% and a precision rate of 86.07\%. As the cryptocurrency market is heavily dependent on Bitcoin, it may also be incorporated into a trading strategy and anticipate not only the return of the Bitcoin price but also market variety and fluctuations. However, in this instance, it is employed for research and study reasons that aid in understanding the variance of cryptocurrencies and utilising machine learning models to forecast returns. Additionally, as discussed in the previous section, a method was used to combine all the models and use a voting strategy to decide the prediction. This method had some efficiency and produced predictions with an accuracy of about 53\%, but it still requires additional work and improvements to be converted into an ensemble learning approach.

XII. Methodology Discussion:

As discussed in the previously, the machine learning development method was used to ensure that this project adheres to the standards for machine learning projects. The methodology offers a structured approach to follow along with that allows flexibility in solving this machine learning problem and is based on these axes: Data collection, Data engineering, Model training, Model optimization, and finally, Model integration, which results in a machine learning project capable of making analytical decisions. Although working on a standard software development project requires a certain way to be able to change, improve, or even delete a part of a project, in this situation, the process allows flexibility in the work by isolating each central element even though it is connected to a next or previous one. Any action can be taken in this manner without endangering other parts (in some cases, a further change must be done with other sections depending on the type of changes). This advantage was obvious when concentrating on the model training and model optimization part when returning to these specific parts to apply any further changes; any change that was applied does not require reviewing the entire project from the start or at a certain point. This benefit was made possible by the methodology that was used. This methodology provides numerous advantages and may be maintained in further work by updating datasets, models, or methodologies.

XIII. Conclusion

This study's primary objective was to determine whether machine learning models could anticipate specific cryptocurrency returns using classification or regression models to predict whether the price will move “Up" or “Down" on that particular day. The methodology utilised in this research was systematic, beginning with the data collection from the Yahoo Finance API, then producing two datasets, one for classification (categorical, which showed two classes “1" implies the price will climb, “0" means the price will decline). The other was for regression, which used numbers to show the daily percent change in price, ranging from -1 to 1. Both datasets were built using the daily close price of Bitcoin over eight years of data. Based on 21 timesteps, time series were created as the main feature of the prediction. Along with the customised ANN model, the initial selection included SVM, XGBoost Classifier, and RandomForest Classifier. Following training, it was determined that RandomForest outperformed all other classifiers listed and then XGBoost with has a less acuuracy than the RandomForest and more balanced score, especially precision. The major regression indicators for each model were highly similar, making it difficult to determine how well the model would perform. As a result, the evaluation of the regression models was a little different for each chosen model. As a result, a created plot, which compares anticipated price percent changes with original values, served as the basis for this comparison. It reveals that only the XGBoost Regressor outperformed all evaluated models, including LSTM, the model that is best suited for time series projects. The outputs of the regression models were converted into a categorical label with "0" and "1" classes, and an additional method was devised to calculate accuracy to compare scores from regression and classification since the project is accuracy-focused. The classification is better at foretelling whether the price will move "Up" or "Down" on a particular day.

Most of the selected models can be employed in various profitable, analysis or even related studies projects, for example, trading strategy, trading bots, analysis dashboard, and market studies. The initial voting strategy was utilised to avoid the forecast being made by a single decision maker, which opened the door for additional work that can be implemented and enhanced based on the voting approach to predict cryptocurrency market returns using ensemble machine learning. There is already a pre-built machine learning method with the same principle but only supports Classification alone or Regression named: (VotingClassifier, VotingRegressor by Sickit-Learn) this method can be more effective. However, it requires additional research and enhancements. More features are required for this subject as it has been demonstrated is that time series alone are insufficient to achieve higher results than those obtained. One aspect of the additional work that can be done is to expand research for other features that can interact directly with the cryptocurrency market prices, which can be related to finance. This research's one major weakness was the lack of features when it came to building the dataset. In conclusion, this project showed that predicting cryptocurrency returns using classification and regression machine learning models is feasible and could be a useful tool for many other fields or projects since the cryptocurrency market is regarded as a new market.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Jazan University and the Government of the Kingdom of Saudi Arabia for the full sponsoring this research.

REFERENCES

[1] R. Grinberg, “Bitcoin: An Innovative Alternative Digital Currency,” Hastings Sci. & Tech. Law J., vol. 4, p. 160, 2011.

[2] F. Mai, Z. Shan, Q. Bai, X. Wang, and R. H. L. Chiang, “How Does Social Media Impact Bitcoin Value? A Test of the Silent Majority Hypothesis,” J. Manag. Inf. Syst., vol. 35, no. 1, pp. 19–52, 2018.

[3] S. Sapuric and A. Kokkinaki, “Bitcoin Is Volatile! Isn’t That Right?,” in International Conference on Business Information Systems, 2014, pp. 255–265.

[4] M. Polasik, A. I. Piotrowska, T. P. Wisniewski, R. Kotkowski, and G. Lightfoot, “Price Fluctuations and the Use of Bitcoin: An Empirical Inquiry,” Int. J. Electron. Commer., vol. 20, no. 1, pp. 9–49, 2015.

[5] S. Bhambhwani, S. Delikouras, G. M. Korniotis, et al., Do Fundamentals Drive Cryptocurrency Prices?, Centre for Economic Policy Research, 2019.

[6] G. Aggarwal, V. Patel, G. Varshney, and K. Oostman, “Understanding the Social Factors Affecting the Cryptocurrency Market,” arXiv preprint arXiv:1901.06245, 2019.

[7] D. Hobson, “What Is Bitcoin?,” XRDS: Crossroads, ACM Mag. Students, vol. 20, no. 1, pp. 40–44, 2013.

[8] P. Ciaian, M. Rajcaniova, et al., “Virtual Relationships: Short- and Long-Run Evidence from Bitcoin and Altcoin Markets,” J. Int. Financ. Mark. Inst. Money, vol. 52, pp. 173–195, 2018.

[9] Kumar, A. S., & Ajaz, T. (2019). Co-movement in crypto-currency markets: evidences from wavelet analysis. Financial Innovation, 5, 26.

[10] R. Lukyanenko, A. Castellanos, J. Parsons, M. Chiarini Tremblay, and V. C. Storey, “Using Conceptual Modeling to Support Machine Learning,” in International Conference on Advanced Information Systems Engineering, 2019, pp. 170–181.

[11] W. Maass and V. C. Storey, “Pairing Conceptual Modeling with Machine Learning,” Data & Knowledge Engineering, vol. 134, p. 101909, 2021.