Data Driven Level Set Fuzzy Modeling for Cryptocurrencies Price Forecasting

Leandro Maciel¹, Rosangela Ballini² and Fernando Gomide³

¹Department of Business Administration, Faculty of Economics, Business and Accounting, University of São Paulo, São Paulo, Brazil
²Department of Economic Theory, Institute of Economics, University of Campinas, São Paulo, Brazil
³Department of Computer Engineering and Automation, School of Electrical and Computer Engineering, University of Campinas, São Paulo, Brazil

Keywords: Data Driven Fuzzy Modeling, Cryptocurrency, Forecasting.

Abstract: The paper develops a data-driven fuzzy modeling procedure based on level set to forecast cryptocurrencies prices. Data-driven level set is a novel fuzzy modeling method that differs from linguistic and functional fuzzy models in how the fuzzy rules are built and processed. The level set-based model outputs the weighted average of output functions associated with the fuzzy rules. Output functions map the activation levels of the fuzzy rules directly in the model outputs. Computational experiments are done to evaluate the level set method to forecast the closing prices of Bitcoin, Ethereum, Litecoin and Ripple. Comparisons are made with ARIMA, ETS, MLP and naïve random walk. The results suggest that the random walk outperforms most methods addressed in this paper, but it is surpassed by the level set model for Ethereum. When performance is measured by the direction of price change, the level set-based fuzzy modeling performs best amongst the remaining methods.

1 INTRODUCTION

Since the creation of Bitcoin (Nakamoto, 2008), one of the most popular cryptocurrencies, a rapid growth of the digital coin market has been verified. According to CoinMarketCap website, the market capitalization of cryptocurrencies on February 14, 2022 was higher than USD 1.91 trillion, with Bitcoin accounting for more than USD 810 billion¹. In contrast to traditional cash systems, advantages of cryptocurrencies are decentralization, security and privacy, easy transfer of funds, and lower transaction costs (Mukhopadhyay et al., 2016). On 2022, more than 2,000 types of cryptocurrencies are available for public trading, which reveals the significance of digital coins as an electronic payment system also as a financial asset, attracting substantial interest from the general public, investors, and researchers (Zhang et al., 2021; Balciilar et al., 2017).

One particular feature of most cryptocurrencies is the high volatile price dynamic, which directly affects investors and speculators profits and losses. The low correlations of digital coins with conventional assets also make the analysis of future price fluctuations more difficult (Parfenov, 2022). Besides being a high complex and risky market, cryptocurrencies still represent an alternative investment instrument, providing an alternate for portfolio diversification (Sun et al., 2020). For example, (Selmi et al., 2018) stated that Bitcoin serves as a hedge, a safe haven, and a diversifier for oil price movements in terms of diversification opportunities and downside risk reductions. Hence, the development of accurate price-forecasting models for cryptocurrencies is of key interest by market participants.

The price of cryptocurrencies is influenced by many factors that induce volatility such as the movements of macroeconomic variables, news and fake news, government policies, and social media contents (Philippas et al., 2019). Researchers have developed many models to predict future prices movements of cryptocurrencies, which can be broadly or-
organized into: i) traditional time-series model such as Autoregressive Integrated Moving Average (ARIMA) (Tandon et al., 2021), and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models (Fung et al., 2021); and ii) machine learning approaches (Chowdhury et al., 2020), such as Support Vector Machines (SVM) (Hitam et al., 2019), neural networks (Zhang et al., 2021), and deep neural nets (Lahmiri and Bekiros, 2019).

Few studies have addressed price forecasting of cryptocurrencies using fuzzy models. For instance, (Atsalakis et al., 2019) developed a hybrid neuro-fuzzy model to forecast the direction in the change of the daily price of Bitcoin. The method outperformed two other computational intelligence models, a simple neuro-fuzzy based approach, and an artificial neural network. Also considering Bitcoin price forecasting, (Garcia et al., 2019) used an evolving granular fuzzy-rule-based model which has a modified rule structure that includes reduced-term consequent polynomials, supplied with an incremental learning algorithm that simultaneously inputs missing data, and updates model parameters and structure. The authors indicated the high accuracy of the suggested approach when compared with fuzzy and neuro-fuzzy evolving modeling methods for Bitcoin price prediction.

This paper develops a data-driven fuzzy model to forecast time series of cryptocurrencies. Data-driven fuzzy modeling uses the concept of level sets to shape a novel fuzzy modeling paradigm. It differs from previous fuzzy modeling paradigms in the way that the fuzzy rules are built and processed. The level set method outputs the weighted average of output functions, functions that map the activation levels of the rules in points of the output variable domain. It has been shown that the data driven level set method (LSM) is simple, effective, and transparent (Maciel et al., 2022). The efficacy of LSM modeling is evaluated in forecasting the daily closing prices of the four most traded cryptocurrencies: Bitcoin, Ethereum, Litecoin and Ripple. Its performance is compared with the autoregressive integrated moving average (ARIMA), the exponential smoothing state model (ETS), the naive random walk, and a multi-layer neural network (MLP) benchmarks.

Since cryptocurrencies price forecasting is still an open topic in the literature, due to the particularities of the market, fuzzy techniques appear as a potential modeling tool, mainly when dealing with the imprecision of the digital coin price movements. For example, in March 2021, Elon Musk had announced in a series of tweets that people could begin buying Tesla cars with Bitcoin, after which the prices of Bitcoin rose about 5% afterwards\(^2\). The data driven level set method is a potential candidate to model and forecast nonlinear and time-varying dynamics such as cryptocurrency prices. The LSM also seems appropriate to model time series that are affected by intangibles like market sentiments which, because its fuzzy nature, limits the expressiveness of statistical and machine techniques.

After this introduction, the paper proceeds as follows. Section 2 summarizes the data driven fuzzy modeling based on the level set framework. Computational experiments concerning modeling and price forecast of cryptocurrencies are reported in Section 3. Finally, Section 4 concludes the paper and lists topics for future development.

2 DATA DRIVEN LEVEL SET MODELING

This section summarizes the data driven fuzzy modeling based on the notion of level set. A detailed coverage is given in (Leite et al., 2022). Consider a fuzzy model whose fuzzy rules are as follows

\[ R_i: \text{if } A_i(x) \text{ then } y \in B_i \]  

where \( i = 1, 2, \ldots, N \) and \( A_i \) and \( B_i \) are convex fuzzy sets with membership functions \( A_i(x) : \mathcal{X} \rightarrow [0, 1] \) and \( B_i(y) : \mathcal{Y} \rightarrow [0, 1] \). Given an input \( x \in \mathcal{X} \), the data-driven level set method is as follows (Leite et al., 2022; Yager, 1991).

1. Compute the activation degree of each rule \( R_i \) as

\[ \tau_i = A_i(x) \]  

2. Find the level set \( B_{\tau_i} \) for each \( \tau_i \)

\[ B_{\tau_i} = \{ y \mid \tau_i \leq B_i(y) \} = [y_{il}, y_{iu}] \]  

3. Compute the midpoint of the level set

\[ m_i = \frac{y_{il} + y_{iu}}{2} \]  

4. Compute the model output \( y \) as

\[ y = \frac{\sum_{i=1}^{N} \tau_i m_i}{\sum_{i=1}^{N} \tau_i} \]  

where \( y_{il} \) is the lower bound, and \( y_{iu} \) is the upper bound of the level set. When fuzzy set \( B_i \) is discrete, \( m_i \) is the average of the elements of \( B_i \). We assume

that there exists an \( i \) such that \( \tau_i > 0 \), otherwise the output is made null.

Let \( F_i(\tau_i) = m_i(\tau_i) \), and \( D = \{ (x^k, y^k) \}, x^k \in R^d \), \( y^k \in R \) such that \( y^k = f(x^k), k = 1, 2, \ldots, K \) be a data set. The goal is to build a fuzzy model \( \mathcal{F} \) to approximate the function \( f \) using \( D \) where

\[
\mathcal{F}(\tau) = \frac{\sum_{i=1}^{N} \tau_i F_i(\tau_i)}{\sum_{i=1}^{N} \tau_i} \tag{6}
\]

In the simplest case \( \mathcal{F} \) is affine

\[
F_i(\tau_i) = v_i \tau_i + w_i \tag{7}
\]

Coefficients \( v_i \) and \( w_i \) can be estimated using least squares-based procedures, regularized, recursive, or alternative solutions. Here we use the pseudo inverse-based solution. The essential steps are as follows.

For each data pair \( (x^k, y^k) \) compute the activation degrees \( \tau_i^k = A_i(x^k), i = 1, 2, \ldots, N \), and let \( e^k = \sum_{i=1}^{N} \tau_i^k \). From (6) and (7), the corresponding output is

\[
\zeta^k = \frac{\sum_{i=1}^{N} \tau_i^k (v_i \tau_i + w_i)}{e^k} + \cdots + \frac{\tau_N (v_N \tau_N + w_N)}{e^k} \tag{8}
\]

Let \( d^k = [\tau_1^k / s^k, \tau_2^k / s^k, \ldots, \tau_N^k / s^k] \) and let the vector of parameters \( u = [v_1, w_1, \ldots, v_N, w_N]^T \). The expression (8) becomes

\[
\zeta^k = d^k \cdot u, \quad k = 1, \ldots, K \tag{9}
\]

If we let \( z = [z_1, \ldots, z_K]^T \), \( D = [d_1^T, \ldots, d_K^T]^T \), and \( y = [y_1, \ldots, y_K]^T \), then the set of equations (9) can be expressed compactly as \( z = Du \). The vector of parameters \( u \) is the solution of \( \min_u ||y - z||^2 \), namely

\[
u = D^+ z \tag{10}
\]

where \( D^+ \) is the Moore-Penrose pseudo inverse of \( D \) (Serre, 2010). Recalling that \( d = d(\tau_1, \ldots, \tau_N) \) and that \( \tau_i = A_i(x) \), the model output for input \( x \) is

\[
y = d \cdot u \tag{11}
\]

which is equivalent to (5).

In sum, the data driven level set method (LSM) is as follows:

1. Cluster the data set \( D \) into \( N \) clusters.
2. Assign membership function \( A_i \) to cluster \( i = 1, \ldots, N \).
3. Find consequent vector of coefficients using (10).

As it is well known in the fuzzy modeling literature, clustering can be done to identify the \( A_i \)'s using any clustering algorithm such as the fuzzy c-means (FCM) or its variations, adaptive vector quantization, grid, or knowledge-based granulation. The membership function parameters can be tuned using context knowledge, or an appropriate search procedure.

### 3 COMPUTATIONAL RESULTS

#### 3.1 Data

This section reports the computational experiments using datasets of the daily closing prices of four leading cryptocurrencies: Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC) and Ripple (XRP)\(^3\). The data was divided in in-sample and out-of-sample sets, for models training and testing, respectively. Table 1 summarizes the in-sample and out-of-sample sets for each cryptocurrency. In the table \( T \) denotes the number of samples (size) of the data set. The data cover periods of high volatility for each cryptocurrency. The idea is to avoid time periods with flat or nearly stable price values because, in these situations, prediction is easier, which explains why the sets are different for each digital coin. Forecasting is one-step-ahead. The modeling and forecasting techniques considered for comparison are the autoregressive moving average (ARIMA) (Box et al., 2016), the exponential smoothing state space model (ETS) (Hyndman et al., 2008), the naïve random walk (RW), and a multilayer perceptron neural network (Haykin, 2009).

<table>
<thead>
<tr>
<th>Crypto</th>
<th>In-Sample</th>
<th>Out-of-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Start</td>
<td>End</td>
</tr>
<tr>
<td>BTC</td>
<td>01/01/2021</td>
<td>07/16/2021</td>
</tr>
<tr>
<td>ETH</td>
<td>01/01/2021</td>
<td>07/16/2021</td>
</tr>
<tr>
<td>LTC</td>
<td>01/01/2021</td>
<td>07/16/2021</td>
</tr>
<tr>
<td>XRP</td>
<td>01/01/2021</td>
<td>07/16/2021</td>
</tr>
</tbody>
</table>

Performance evaluation of methods is done using the mean squared error (MSE), the root mean squared error (RMSE), and the normalized RMSE (NRMSE), respectively:

\[
\text{MSE} = \frac{1}{T} \sum_{t=1}^{T} (y_t - \hat{y}_t)^2 \tag{12}
\]

\[
\text{RMSE} = \sqrt{\text{MSE}} \tag{13}
\]

\[
\text{NRMSE} = \frac{\text{RMSE}}{sd(y_{test})} \tag{14}
\]

where \( y_t \) is the actual price at time \( t \), \( \hat{y}_t \) the forecasted price at \( t \), \( T \) is the size of the out-of-sample data set,

---

\(^3\)Selection done choosing cryptocurrencies with the highest liquidity and market capitalization. Data source: https://coinmarketcap.com/
sd(·) is the standard deviation, and $y_{test}$ the actual out-of-sample data.

Additionally, in practice the direction of price change is as important as, sometimes even more important than the magnitude of the forecasting error (Burns and Moosa, 2015). A measure of forecast direction is:

$$DA = \frac{1}{T} \sum_{t=1}^{T} Z_t, \quad Z_t = \begin{cases} 1, & \text{if } (\hat{y}_{t+1} - y_t)(y_{t+1} - y_t) > 0, \\ 0, & \text{otherwise}. \end{cases}$$

(15)

Naïve random walk uses the current price as a forecast of the next time step:

$$\hat{y}_t = y_{t-1}.$$  

(16)

RW forecasting follows the weak form of market efficiency, which states that future securities prices are random and are not affected by past events. It assumes that information of stock prices are reflected in the current prices and has no relationship with the past market prices. Clearly, RW is not capable to predict price direction.

As in time series modeling and forecast methodology, MLP and LSM model and forecast cryptocurrency prices using lagged values of the series as follows:

$$\hat{y}_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-d}),$$

(17)

where $d$ is the number of lagged closing prices.

### 3.2 Results

Evaluation and comparison of LSM with ARIMA, ETS, RW and MLP in one-step-ahead cryptocurrencies closing prices forecasting uses out-of-sample data as testing data. LSM was implemented in Python, and ARIMA, ETS and MLP were constructed using R forecast package. Table 2 shows the parameter structure of each method. In ARIMA(p,d,q), $p$, $d$ and $q$ stand for the number of autoregressive, difference, and moving average terms, respectively. In ETS$(a,dr,sea)$ $e$, $r$, and $s$ mean error, trend, and seasonal type, respectively. For each of these $A$, $M$ and $N$ denote additive, multiplicative, and none, respectively. MLP($d,h_1,h_2,\ldots,h_n$) denotes a neural network with as many inputs as the number lagged values of the series model $d$, and with $h_i$ neurons in the $i$-th hidden layer, $i = 1, \ldots, n$. The MLP neural networks use sigmoid activation functions in the hidden layer with linear output layer, trained with backpropagation. The structures of ARIMA, ETS and MLP are selected automatically by R to produce the highest accuracy in the in-sample set. LSM($d,N$) means a LSM model with $d$ lagged values, and $N$ fuzzy rules with Gaussian membership functions. LSM structures were chosen experimentally and the simplest model with best accuracy was chosen. In this paper, LSM used Gaussian membership functions. As the model output functions are affine, the parameters are the Gaussians modal values and dispersions, and output functions coefficients (a total of four parameters for each rule). Therefore, according to the structure of the models given in Table 2, the LSM has 8, 8, 12 and 12 parameters for BTC, ETH, LTC and XRP, respectively. Additionally, LSM used fuzzy c-means clustering for data partitioning.

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE</th>
<th>RMSE</th>
<th>NRMSE</th>
<th>DA (%)</th>
</tr>
</thead>
</table>
| Panel A: BTC
| RW     | 2.78961 | 1.67021 | 0.19697 | -      |
| ARIMA  | 2.93753 | 1.71392 | 0.20212 | 47.19  |
| ETS    | 2.84838 | 1.68771 | 0.19903 | 48.77  |
| MLP    | 3.25329 | 1.80369 | 0.21271 | 47.00  |
| LSM    | 2.82390 | 1.68040 | 0.19860 | 52.08  |
| Panel B: ETH
| RW     | 2.16047 | 1.46985 | 0.21153 | -      |
| ARIMA  | 2.36282 | 1.53715 | 0.22121 | 49.02  |
| ETS    | 2.50985 | 1.58425 | 0.22799 | 47.30  |
| MLP    | 2.38548 | 1.54450 | 0.22227 | 48.09  |
| LSM    | 2.34059 | 1.52990 | 0.22017 | 51.04  |
| Panel C: LTC
| RW     | 0.82744 | 0.99664 | 0.28183 | -      |
| ARIMA  | 0.87168 | 0.93364 | 0.28926 | 50.20  |
| ETS    | 0.86557 | 0.93036 | 0.28825 | 50.87  |
| MLP    | 0.97365 | 0.98674 | 0.30571 | 49.16  |
| LSM    | 0.83820 | 0.91550 | 0.28440 | 56.38  |

Panel D: XRP
| RW     | 0.00244 | 0.04941 | 0.25065 | -      |
| ARIMA  | 0.00244 | 0.04940 | 0.25062 | 51.66  |
| ETS    | 0.00244 | 0.04944 | 0.25082 | 51.74  |
| MLP    | 0.00263 | 0.05126 | 0.26007 | 49.89  |
| LSM    | 0.00244 | 0.04940 | 0.25061 | 53.72  |

Naïve random walk outperforms all competitors from the point of view of MSE, RMSE and NRMSE metrics, except for Ripple (XRP) when LSM gives the best result amongst the remaining ones (see Table 3). These results are consistent with the “Meese-Rogoff
“puzzle” (Meese and Rogoff, 1983) which states that exchange rate forecasting models do not outperform random walk. Interestingly, our simulations show that this is also the case for cryptocurrency forecasts. For BTC, ETH and LTC, LSM reaches the highest accuracy after RW. ARIMA, ETS and MLP achieve similar performance, but LSM forecasts are either the closest, or better than those of RW.

The literature (Moosa and Burns, 2014) and (Burns and Moosa, 2015) reports empirical findings showing that forecasting models can outperform the naïve random walk for out-of-sample data if the performance is measured by economic/financial metrics such as the direction of change and/or profitability as in forecast-based trading operations. Indeed, this is also the case in cryptocurrency price forecasting. As Table 3 shows, LSM outperforms ARIMA, ETS, and MLP from the point of view of the direction accuracy (DA). The random walk underperforms all the methods when the direction is used for comparison. This is of utmost importance when trading strategies use direction, because the potential to anticipate price change is crucial for the success.

To further illustrate the efficiency of LSM in cryptocurrencies forecast, Figures 1–4 show the actual closing prices and the corresponding LSM forecasts developed for BTC, ETH, LTC and XRP using test data, respectively. The figures reveals that LSM accurately predict price dynamics of the digital coins considered in this work, and unfolds as a potential forecast tool to develop trading strategies in cryptocurrency market.

4 CONCLUSION

This paper has developed a data-driven fuzzy level set-based model to forecast cryptocurrencies prices. It differs from alternative fuzzy models mainly in the way that the fuzzy rules are built and processed. A level set-based model outputs the weighted average of rule output functions, functions that map the activation levels of the rules directly in a model output. Computational experiments were done for one-step-ahead forecasting of closing prices of Bitcoin, Ethereum, Litecoin and Ripple. Comparison of the data driven level set-based method was done against ARIMA, ETS, MLP and naïve random walk. The results indicate that random walk outperforms all the competitors addressed in this paper, except for Ripple, where LSM produced the highest out-of-sample accuracy. However, when performance is measured by the direction of price change, the level set-based fuzzy modeling rank is the highest. Future work shall consider the use and evaluation of the data-driven level set method in trading strategies to measure the forecast performance in profitability terms.
ACKNOWLEDGEMENTS

This work was supported by the Brazilian National Council for Scientific and Technological Development (CNPq) under grants 304456/2020-9, 04274/2019-4 and 302467/2019-0, by the Ripple Impact Fund, a donor advised fund of the Silicon Valley Community Foundation, grant 2018-196450(5855) as part of the University Blockchain Research Initiative, UBRI, and by the Research Foundation of the State of São Paulo (FAPESP), grant 2020/09838-0.

REFERENCES


