IEEE WCCI 2020

IEEE World Congress on Computational Intelligence

Virtual Conference – July 19-24, 2020

Multi-step ahead Bitcoin Price Forecasting Based on VMD and Ensemble Learning Methods

R.G. da Silva, M.H.D.M. Ribeiro, N. Fraccanabbia, V.C. Mariani, L.S. Coelho

Industrial & Systems Engineering Graduate Program (PPGEPS).
Pontifical Catholic University of Parana (PUCPR), Parana, Brazil

Agenda

- Introduction
- Dataset
- Methodology
- Results
- Conclusion
- Acknowledgments

Introduction

- A cryptocurrency is a digital asset designed to work as a exchange that uses strong cryptography to secure financial transactions.
- Bitcoin is considered the first decentralized cryptocurrency, and it was first released as open-source software in 2009.
- With the emergence of the cryptocurrency market, the Bitcoin, its leading currency, has captured global attention.

Introduction

Bitcoin price time series has a high volatility

Forecasting Bitcoin price as accurate as possible is a challenge

Hybrid ensemble learning models can handle this volatility

Variational mode decomposition (VMD) and Stacking-ensemble learning (STACK)

Introduction

Objective

- To develop a heterogeneous stacking-ensemble learning model for Bitcoin price forecasting multi-step ahead (one, two and three days ahead).
- The proposed model is composed by VMD and STACK approaches, and heterogeneous forecasting models.

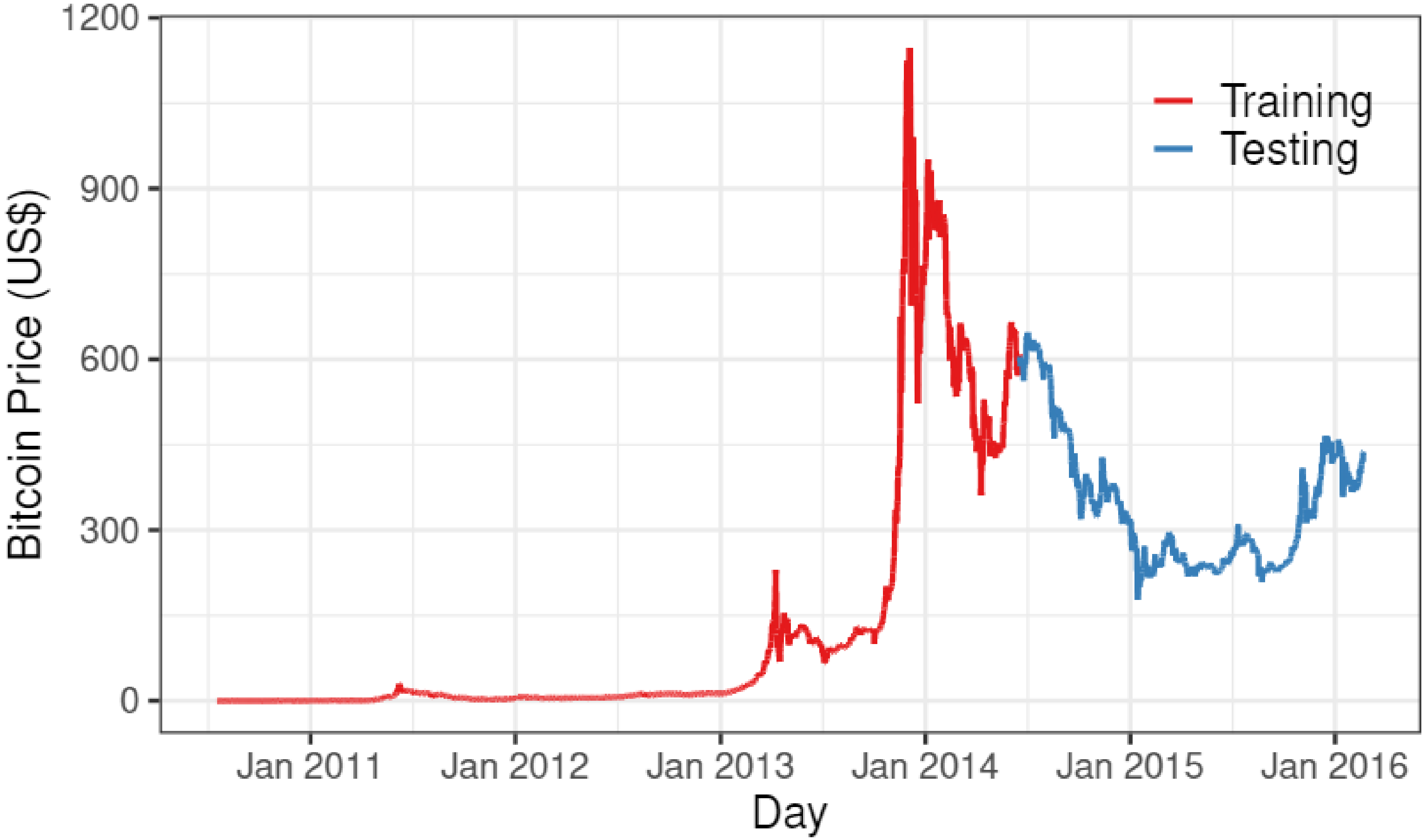
Dataset

- Daily observations from **July 28th 2010** to **February 21st 2016**

- Observations number: **2045**

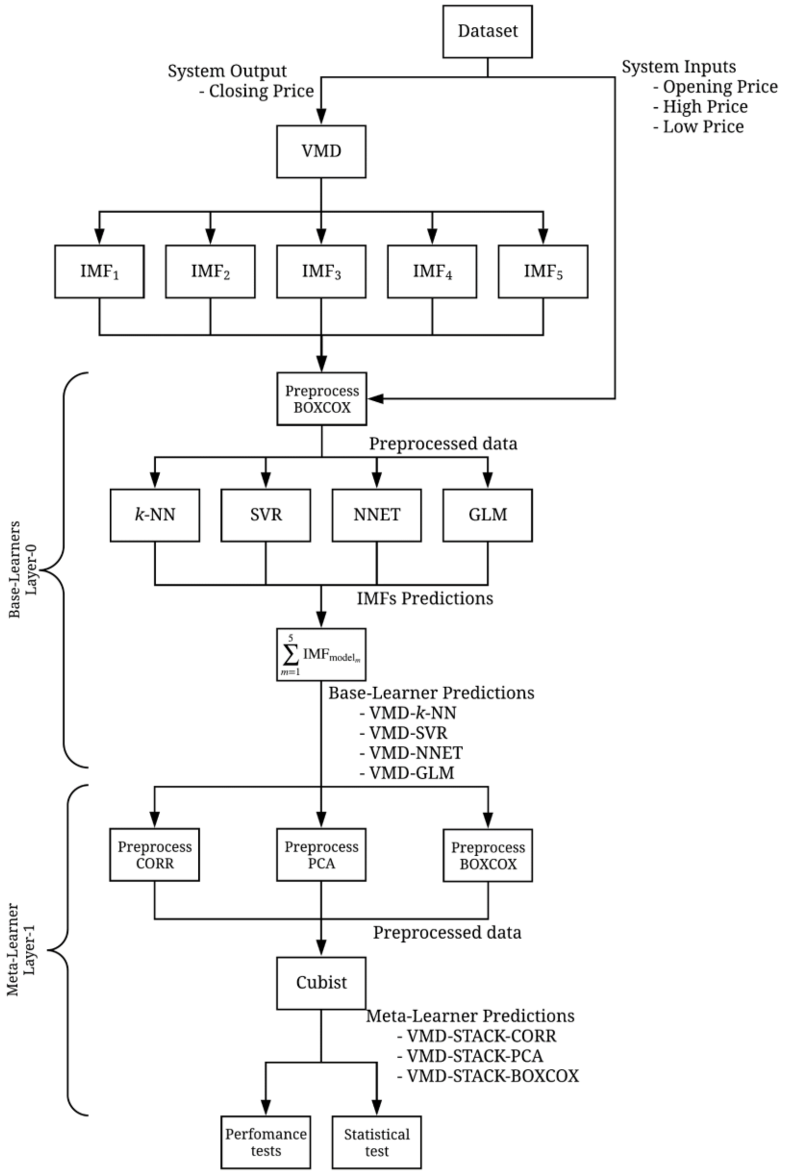
TABLE I
INPUTS AND OUTPUT OF THE SYSTEM

Type	Description	Unit Measure
Input	Opening Price	US$
Input	High Price	
Input	Low Price	
Output	Closing Price	



Methodology

- **VMD** decomposed the data into 5 components;
- A **Box-Cox transformation** as preprocess;
- Four different algorithms (models) as base-learners:
    - ***k*-NN** (*k*-Nearest Neighbor)
    - **SVR** (Support Vector Regression with Linear kernel)
    - **NNET** (Feed-forward Neural Network)
    - **GLM** (Generalized Linear Model)
- Base-learners predictions were preprocessed by:
    - **CORR** (Correlation Matrix)
    - **PCA** (Principal Component Analysis)
    - **Box-Cox transformation**
- Stacking meta-learner:
    - **Cubist Regression**



Methodology

- **Performance measures:**

$$\mathrm{RRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{\frac{1}{n}\sum_{i=1}^{n} y_i}, \quad \mathrm{sMAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{(|y_i| + |\hat{y}_i|/2)}\right|, \quad \mathrm{APE} = \frac{|y_i - \hat{y}_i|}{y_i}$$

- **Statistical test:**

$$\mathrm{DM} = \frac{\frac{\sum_{n}^{i=1}[d_i]}{n}}{\sqrt{\frac{\mathrm{var}(d_i)}{n-1}}}$$

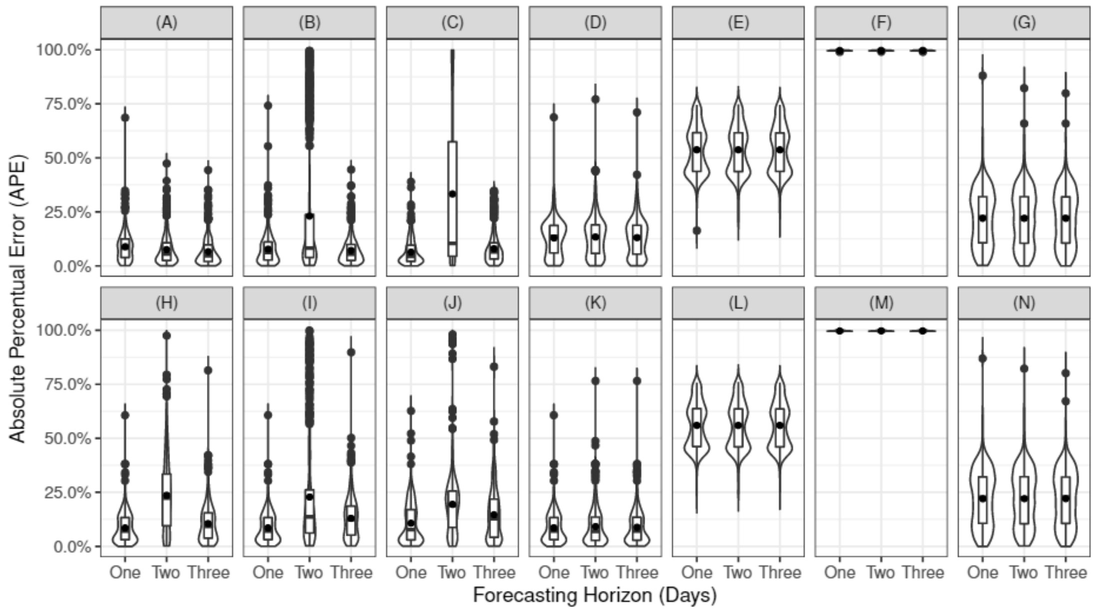
Diebold-Mariano test

Results

TABLE IV
PERFORMANCE MEASURES RESULTS OF THE MODELS

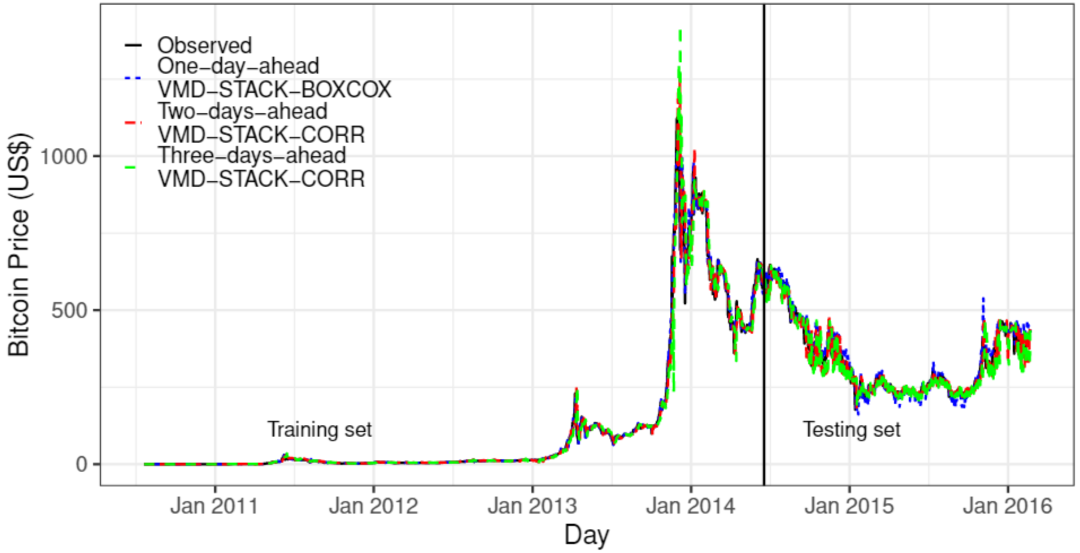
Model	Forecasting Horizon					
	One-day-ahead		Two-days-ahead		Three-days-ahead	
	sMAPE	RRMSE	sMAPE	RRMSE	sMAPE	RRMSE
(A) VMD–STACK–CORR*	0.0835	0.0936	**0.0660**	0.0943	**0.0762**	0.1060
(B) VMD–STACK–PCA*	0.0735	0.0915	0.0708	**0.0934**	0.2310	0.3812
(C) VMD–STACK–BOXCOX*	**0.0626**	**0.0800**	0.0836	0.1003	0.5766	0.7108
(D) VMD–k-NN**	0.1189	0.1244	0.1201	0.1274	0.1224	0.1304
(E) VMD–SVR**	0.7506	1.5222	0.7499	1.5187	0.7497	1.5184
(F) VMD–NNET**	1.9870	355.3495	1.9870	354.8142	1.9870	354.7863
(G) VMD–GLM**	0.2153	0.2867	0.2151	0.2858	0.2151	0.2858
(H) STACK–CORR*	0.0804	0.0935	0.0976	0.1099	0.2033	0.1954
(I) STACK–PCA*	0.0806	0.0936	0.1185	0.1263	0.2169	0.3395
(J) STACK–BOXCOX*	0.0985	0.1131	0.1315	0.1596	0.3608	0.5100
(K) k-NN**	0.0804	0.0935	0.0824	0.0985	0.0850	**0.1028**
(L) SVR**	0.7925	1.6495	0.7917	1.6458	0.7916	1.6456
(M) NNET**	1.9871	359.3776	1.9871	358.8427	1.9871	358.8427
(N) GLM**	0.2154	0.2863	0.2152	0.2854	0.2152	0.2854

Note: *Cubist as meta-learner; **BoxCox as pre-processing.

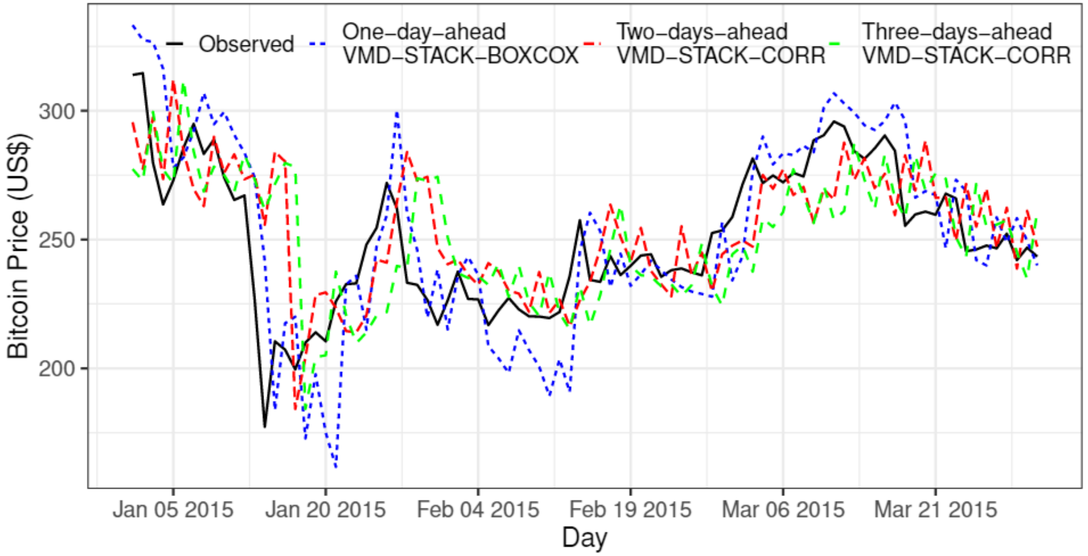


Results

Predicitions for whole dataset



Samples from Jan. 1st 2015 to Mar. 31st 2015



Results

TABLE V
DIEBOLD-MARIANO TEST RESULTS

Model	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)	(N)
							One-day-ahead							
(A)	-	-1.1808	3.1101*	-7.3984*	-12.0835*	-16.5659*	-8.3376*	1.8446	1.8281	-3.7099*	1.8446	-12.2978*	-16.5617*	-8.3427*
(B)	-	-	3.0229*	-5.8826*	-12.0814*	-16.5654*	-8.3103*	1.9772**	1.9679**	-1.6512	1.9772**	-12.2958*	-16.5612*	-8.3153*
(C)	-	-	-	-8.0185*	-12.0883*	-16.5674*	-8.3892*	-0.9842	-0.9961	-4.9253*	-0.9842	-12.3022*	-16.5632*	-8.3948*
							Two-days-ahead							
(A)	-	0.2929	-0.2391	-5.5822*	-6.9657*	-9.5660*	-4.8075*	-2.5678**	-4.5894*	-6.8380*	-0.9036	-7.0893*	-9.5635*	-4.8115*
(B)	-	-	-0.4671	-5.6608*	-6.9654*	-9.5658*	-4.8066*	-2.8868*	-4.9356*	-6.9963*	-1.1455	-7.0891*	-9.5634*	-4.8106*
(C)	-	-	-	-4.6239*	-6.9646*	-9.5655*	-4.7984*	-1.9949**	-3.9650*	-6.4008*	-0.5822	-7.0882*	-9.5631*	-4.8024*
							Three-days-ahead							
(A)	-	-13.8752*	-7.8720*	-3.5211*	-5.4017*	-7.4190*	-3.7026*	-7.0480*	-27.4156*	-13.8618*	0.1871	-5.4978*	-7.4171*	-3.7051*
(B)	-	-	-1.5640	13.9756*	-2.1187**	-6.4100*	10.3861*	13.5846*	5.7783*	-35.6064*	13.8949*	-2.4337**	-6.4087*	10.4037*
(C)	-	-	-	7.8447*	-17.5507*	-7.6499*	8.9298*	7.6728*	4.0853*	-12.5259*	7.8699*	-6.6571*	-7.6474*	8.9214*

Note: *1% significance level; **5% significance level.

Conclusion

- This study proposed a novel heterogeneous decomposition-ensemble learning model by using **VMD** and **STACK** with different preprocessing algorithms to forecast Bitcoin price multi-step-ahead.
- The stacking-ensemble was composed by ***k*-NN**, **SVR**, **NNET** and **GLM**, as base-learners in the first layer, and using **Cubist** as meta-learner in the second layer.
- Indeed, the **VMD-STACK** approach had a **better** performance than compared models in almost all forecasting horizons.
- **For future works**
  - Adopt different combinations of models in both layers of the stacking-ensemble
  - Optimize the hyperparameters of the models
  - Optimize the number of components to be decomposed
  - Use different decomposition methods
  - Increase the forecast horizon interval to more than 3 days ahead

Acknowledgments

CAPES

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

FUNDAÇÃO ARAUCÁRIA
Apoio ao Desenvolvimento Científico e Tecnológico do Paraná

PUCPR

Thank you!