UNC KENAN-FLAGLER BUSINESS SCHOOL

# Application of Machine Learning for Meme Coin Scam Detection

Yingxi Huang, BSBA 2022

Under the direction of Dr. Donghwa Shin

## Abstract

Rising technologies enable people to create, list, and trade cryptocurrencies much easier, cheaper, and faster. However, the low entry barriers and censorship-free environment are abused by scammers. Many new cryptocurrencies—including meme coins created around memes and inside jokes and have no fundamental values—were devised following the launch and success of these technologies. Leveraging pop culture to attract investors, meme coins are especially vulnerable to scams. This thesis establishes meme coin scam detection models with three classification techniques: logistic regression, random forest, and SVM (support vector machine). I find that the ex-ante characteristics of meme coins can help with predicting scams with an accuracy of around 70%. The most effective characteristics are identified. It's also observed that the random forest technique produces overall better performance on validation data compared to the others under this scenario, yet the best performer on test data varies across the three different dependent variables used to categorize scams.

## Introduction

The global market capitalization of cryptocurrency surpassed $1.83 trillion by February 2022 and there were 221 million cryptocurrency users by June 2021. The rising popularity of cryptocurrency and its supporting technology Blockchain incubates both new opportunities and challenges. On the one hand, they protect trade anonymity and transparency and overcome trade barriers such as authority intermediation and transaction cost. Therefore, they have huge potential for positive social changes. It can help with promoting financial inclusion, equality, human rights, and sustainability. On the other hand, as cryptocurrency quickly expands its influences, legislations and regulations fail to keep up, leaving cryptocurrency bringing harm to society. Around $681 million in losses had been caused by major crypto thefts, hacks, and frauds between January to July 2021.

As one of the most popular yet controversial inventions of cryptocurrency, meme coin is on the rise. A meme coin is commonly believed to be a meme-inspired, joke-originated, or pop-culture-related cryptocurrency. Some well-known examples of meme coins are Dogecoin (DOGE) and Shiba Inu (SHIB). From January 2021 to November, the price of Dogecoin has surged by over 8,000% while Shiba Inu gained more than 60,000,000%. Both Dogecoin and Shiba Inu are currently worth more than $20 billion, and they are the two meme coins among the top 10 cryptocurrencies by market capitalization. Unlike traditional cryptocurrencies, which usually have well-designed ecosystems, including their own blockchain, whitepaper, and deflationary policies, most meme coins lack technological and economic foundations. They have no business model, fundamental economics, underlying value, or self-developed supporting technology. They also often have very low prices and short lifecycles, and thus they are often abused by scammers. Decentralized finance (DeFi), decentralized exchanges (DEXes), and Binance Smart Chain are some of the major innovations that powered the rising of meme coin scams such as rug pulls. The literature suggests that social media, especially Reddit, help to exacerbate cryptocurrency scams and thus can be used in analyzing meme coins.

## Materials

This research extracts all independent variables of meme coins from Reedit advertisements because 1. Advertisements reveal how meme coin creators introduce their coins and access the investors. 2. Reddit is one of the most popular online forums and active communities for cryptocurrency. Many scholars have chosen Reddit as their data source for analyzing cryptocurrency.

Prof. Shin and finance Ph.D. student Chuyi Sun at the UNC Kenan-Flagler Business School, collaborating with Prof. Tao Li and Prof. Baolian Wang at the University of Florida Warrington College of Business, have constructed a sample base consisting of 3212 tokens listed on DEXes, which is shared and used in this thesis. By using Python and web crawler, I analyzed the sample's corresponding Reddit advertisement posts and acquired 13 variables that are commonly used in describing meme coins' main features and hypothetically better predictors.

After some explorations, I decided to use the following 3 variables as my dependent variables and the metrics to determine whether a meme coin is a scam or a suspicious coin.

*Dependent Variables*

| Feature | Description | Notation |
|---|---|---|
| Website existence | Whether the official website of the meme coin mentioned in its post still exist | *WebsiteCeased* |
| BSC comment | Whether the meme coin is marked as a scam on BscScan | *BSCScam* |
| Unusual drop in trading volume | Whether there is a drastic decrease in the meme coin's trading volume | *TVUnusualDrop* |

## Methodology

In this study, I use Python and R to form multiple commonly used binary classification algorithms, including logistic regression, random forest, and SVM, to form detection models for meme coin scams. With these models, I identify the most effective predictors from all potential predictors manually selected based on meme coin features collected from 3212 meme coins' Reddit posts. In the end, I will use cross-validation, area under ROC curve, and accuracy to evaluate each model and compare their performances.

Machine Learning for Classification:

1. Logistic Regression: linear model. The second LR prediction model was established with feature selection based on significant p-values.
2. Random Forest: slicing data into rectangles. The second RF prediction model was established with feature selection based on importance level.
3. SVM with radial kernel: using curved lines to separate data

Testing:

1. Ten-fold Cross-Validation: using different segments as training data and test data
2. Accuracy
3. ROC Test: measure the trade-off between FPR and TPR

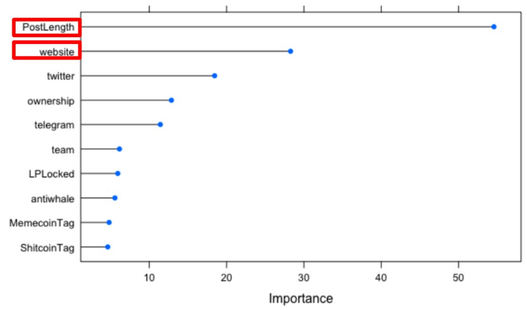
## Results

Effective Predictors:

*Logistic Regression for Website Existence*

| | Estimate | Std. Error | z value | Pr(>\|z\|) | Significance |
|---|---|---|---|---|---|
| **(Intercept)** | 2.50E+00 | 1.76E-01 | 14.186 | < 2e-16 | *** |
| **PostLength** | 2.57E-05 | 1.20E-05 | 2.151 | 0.031502 | * |
| **website** | -1.29E+00 | 1.55E-01 | -8.32 | < 2e-16 | *** |
| **telegram** | -6.45E-01 | 1.75E-01 | -3.678 | 0.000235 | *** |
| **twitter** | -5.90E-01 | 1.30E-01 | -4.549 | 5.39E-06 | *** |
| **ownership** | 7.42E-01 | 1.43E-01 | 5.202 | 1.98E-07 | *** |
| **antiwhale** | 3.44E-01 | 1.68E-01 | 2.051 | 0.040274 | * |

*Importance of Top 10 Independent Variables for Website Existence*




Effective Models

*Comparison of Model Performance on the Validation Data for Website Existence*

| Model | ROC | Accuracy |
|---|---|---|
| logistic regression (all) | 0.7550502 | 0.8106 |
| logistic regression (selected) | 0.7518881 | 0.8125 |
| random forest (all) | 0.8945666 | 0.8164 |
| random forest (selected) | **0.8968361** | **0.8192** |
| SVM with radial kernel (all) | 0.793996 | 0.8153 |
| SVM with radial kernel (selected) | 0.8048585 | 0.8115 |

*Comparison of Model Performance on the Test Data for Website Existence*

| Model | ROC | Accuracy |
|---|---|---|
| logistic regression (all) | 0.7291143 | 0.8037 |
| logistic regression (selected) | 0.7233378 | 0.8048 |
| random forest (all) | 0.7295073 | 0.8037 |
| random forest (selected) | **0.7316936** | 0.8004 |
| SVM with radial kernel (all) | 0.625319 | 0.8082 |
| SVM with radial kernel (selected) | 0.6645107 | **0.8104** |

*Comparison of Model Performance on the Test Data for BSC Comment*

| Model | ROC | Accuracy |
|---|---|---|
| logistic regression (all) | **0.6854757** | **0.8666** |
| logistic regression (selected) | 0.6667989 | 0.8644 |
| random forest (all) | 0.5838141 | 0.8666 |
| random forest (selected) | 0.5881285 | 0.8666 |
| SVM with radial kernel (all) | 0.5536338 | 0.8525 |
| SVM with radial kernel (selected) | 0.5616512 | 0.8525 |

*Comparison of Model Performance on the Test Data for Unusual Drop in Trading Volume*

| Model | ROC | Accuracy |
|---|---|---|
| logistic regression (all) | 0.5126309 | 0.5824 |
| logistic regression (selected) | 0.5286822 | 0.5781 |
| random forest (all) | 0.5209998 | **0.5868** |
| random forest (selected) | 0.5169221 | 0.5857 |
| SVM with radial kernel (all) | 0.523165 | 0.5765 |
| SVM with radial kernel (selected) | **0.5297045** | 0.5765 |

## Conclusion

This thesis achieves the following three main conclusions:

1. Ex-ante characteristics of meme coins can help with predicting scams with an accuracy of around 70%.
2. The most effective predictors suggested by the models are ownership (whether the meme coin renounce or lock ownership), Twitter (whether the meme coin has a Twitter account), website (whether the meme coin has a website), and Telegram (whether the meme coin has a Telegram channel).
3. Random forest overall produces better performance with this data compared with the other two classification algorithms. However, for each dependent variable, the best-performing model varies.

Future studies can use more advanced techniques and various other data sources to predict cryptocurrency scams with improved accuracy and achieve a more in-depth understanding of these scams. More relevant knowledge is needed for regulators to defend justice in the cryptocurrency market, for investors to protect their interests, and for the world to unlock the full potential of cryptocurrency and its related technologies safely and effectively.

## Acknowledgements

This thesis is a valuable learning experience and a milestone for my undergraduate education. I want to express my most genuine gratitude to the following individuals, who helped me a lot to make this thesis possible: Dr. Donghwa Shin, Dr. Yasser Boualam, Dr. Kristin Wilson, Dr. Patricia Harms, Dr. Kim Allen, Mr. Phil Hardy, other Staff Members of the Kenan Scholars Program, my family, and friends.