PUMP AND DUMP DETECTION ON CRYPTO CURRENCIES USING COMPUTER VISION

Jayasundara Abeykoon Waruna Priyankara Wickramasingha

199371P

Degree of Master of Science

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

May 2021

PUMP AND DUMP DETECTION ON CRYPTO CURRENCIES USING COMPUTER VISION

Jayasundara Abeykoon Waruna Priyankara Wickramasingha

199371P

Dissertation submitted in partial fulfilment of the requirements for the Degree Master of Science

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

May 2021

Declaration

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

2021/05/29

J.A.W.P Wickramasingha Date

The above candidate has carried out research for the Master's dissertation under my supervision.

………………………………….. 2021/05/29

Dr. Uthayasanker Thayasivam Date

Abstract

Inspired by the immense success shown by artificial neural networks in computer vision on images classification, we propose a novel framework to detect one of the rife fraudulent financial manipulations in crypto currency trading world known as pump and dump. The representation of crypto currency financial charts was re-imagined ameliorating the classification by taking advantage of some of the very recent advancements of time series to spatial encoding techniques of Gramian Angular Field (GAF), Markov Transition Field (MTF) and Recurrence plots (RP) that are capable of spatially encoding the temporal financial time series data in the form of images. Encoded images were then used to train several convolutional neural network architectures which have been able to achieve a very high precision, recall and F1 values close to 99% over the unseen data for the above classification task. This is one of the first of such researches in pump and dump detection in crypto currencies using computer vision. This approach has the potential to be extended in detecting predefined shapes of time series charts.

Key words

cryptocurrency, pump and dump, imbalanced time series classification, spatial encoding temporal data, Gramian angular field, Markov transition field, recurrence plot, cnn, machine learning, market surveillance, class imbalance problem, synthetic minority class oversampling technique

Acknowledgement

My utmost gratitude and appreciation goes to Mrs. Kaushalya Kularatnam - Head of Quantitative Surveillance & Technology London Stock Exchange in making this project a success by providing the initial idea of this research and facilitating all the business, technical and required cloud resources for this research from the very beginning.

I would like to give my utmost gratitude to my supervisor Dr. Uthayasanker Thayasivam - Senior Lecturer of department of Computer Science and Engineering of University of Moratuwa, giving the outstanding consultation and supervision throughout the project.

I am really grateful to Dr. Rasika Withanawasam - Senior Software Architect at Millennium Surveillance systems, sharing all the expertise and knowledge in financial surveillance world and providing the golden opportunity to conduct this research with the help of London Stock Exchange's (LSE) market supervision team.

I would like to extend my gratitude to all the wonderful people from London Stock Exchange Quantitative Surveillance team and LSEG Technology who supported verifying the classifications by these models and assisted me expand the knowledge and exposure to capital market and market surveillance domains.

Last but not least, I would like to extend my gratitude to my beloved wife Ranmali and beloved baby Kiara who were behind me all the time giving me all the strength, courage and dedication for the time required for this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Definition
CNN	Convolutional Neural Network
GAF	Gramian Angular Field
GASF	Gramian Angular Summation Field
GADF	Gramian Angular Difference Field
MTF	Markov Transition Field
RP	Recurrence Plot
PND	Pump and Dump
OHLCV	Open High Low Close Volume
MTS	Multivariate Time Series Sensory
SMOTE	Synthetic Minority Oversampling Technique

LIST OF APPENDICES