

**SENTIMENT-BASED SUPPORT VECTOR
MACHINE OPTIMIZED BY METAHEURISTIC
ALGORITHMS FOR CRYPTOCURRENCY
FORECASTING**

BY

NOR AZIZAH BINTI HITAM

**A thesis submitted in fulfilment of the requirement for the
degree of Doctor of Philosophy in Computer Science**

**Kulliyah of Information and Communication Technology
International Islamic University Malaysia**

AUGUST 2022

ABSTRACT

Time series are used to model a variety of financial phenomena. The cryptocurrency forecasting problem is the focus of this thesis, which investigates time series forecasting challenges in finance. However, earlier research has neglected to consider the importance of sentiment and public opinion in today's market. The Commodity Channel Index (CCI), historical data and a machine learning algorithm are also employed in this study to improve the accuracy of time series forecasting. By employing hyperparameter optimization, this thesis intends to offer a novel sentiment-based support vector machine optimised by particle swarm and moth-flame optimization algorithms (SVMPSOMFO). PSO, GA, WOA, GOA, GWO, HS and MFO are compared against the proposed algorithm's performance for predicting cryptocurrency prices. A thorough investigation and discussion of all experimental results on different datasets are performed. From the findings, SVMPSOMFO outperforms other optimization methods in terms of accuracy rate when compared to a prediction model that excludes sentiment information. In addition, statistical tests are performed to validate the outcomes of the study.

خلاصة البحث

تُستخدم السلاسل الزمنية لنمذجة ظواهر مالية متنوعة. ومشكلة التنبؤ بأسعار العملات المشفرة هي محور هذه الأطروحة، التي تبحث في تحديات التنبؤ بالسلسلة الزمنية في الأمور المالية. ومع ذلك، فقد أهملت الأبحاث السابقة النظر في أهمية المشاعر والرأي العام في السوق اليوم. وقد تم استخدام مؤشر قناة السلع (CCI) والبيانات التاريخية وخوارزمية التعلم الآلي في هذه الدراسة لتحسين دقة تنبؤ السلاسل الزمنية. من خلال استخدام تحسين المعلمات الفائقة، تهدف هذه الأطروحة إلى تقديم آلة المتجه الداعم بصورة جديدة قائمة على المشاعر و محسنة بواسطة خوارزميات تحسين عناصر السرب ولهب العثة (SVMP SOMFO). تم مقارنة PSO و GA و WOA و GOA و GWO و HS و MFO مع أداء الخوارزمية المقترحة للتنبؤ بأسعار العملات المشفرة. كما تم إجراء تحقيق و مناقشة شاملين لجميع النتائج التجريبية على مجموعات البيانات المختلفة. يظهر من النتائج، تفوق SVMP SOMFO على طرق التحسين الأخرى من حيث معدل الدقة عند مقارنته بنموذج التنبؤ الذي يستبعد معلومات المشاعر. بالإضافة إلى ذلك، تم إجراء الاختبارات الإحصائية للتحقق من صحة نتائج الدراسة.

APPROVAL PAGE

The thesis of Nor Azizah binti Hitam has been approved by the following:

Amelia Ritahani binti Ismail
Supervisor

Normaziah binti Abdul Aziz
Co-Supervisor

Rizal Mohd Nor
Internal Examiner

Nazlia Omar
External Examiner

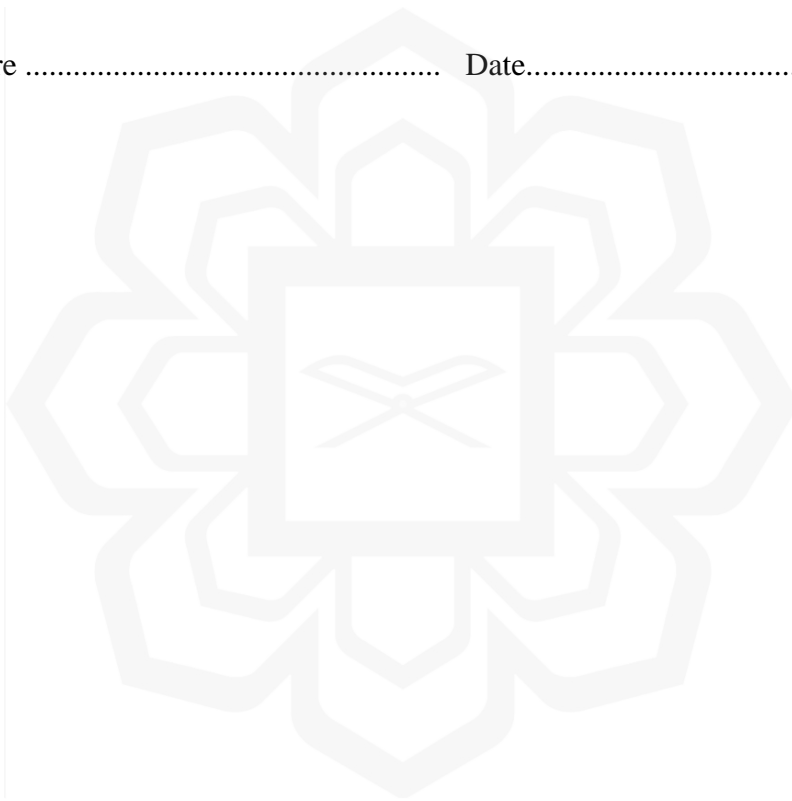
Mohammed Elwathig Saeed Mirghani
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Nor Azizah binti Hitam

Signature Date.....



INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA
DECLARATION OF COPYRIGHT AND AFFIRMATION OF
FAIR USE OF UNPUBLISHED RESEARCH

SENTIMENT-BASED SUPPORT VECTOR MACHINE
OPTIMIZED BY METAHEURISTIC ALGORITHMS FOR
CRYPTOCURRENCY FORECASTING

I declare that the copyright holders of this thesis are jointly owned by the student and IIUM.

Copyright © 2022 International Islamic University Malaysia. All rights reserved.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may be used by others in their writing with due acknowledgement.
2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purposes.
3. The IIUM library will have the right to make, store in a retrieved system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Nor Azizah binti Hitam

.....
Signature

.....
Date

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

**DECLARATION OF COPYRIGHT AND AFFIRMATION OF
FAIR USE OF UNPUBLISHED RESEARCH**

**SENTIMENT-BASED SUPPORT VECTOR MACHINE
OPTIMIZED BY METAHEURISTIC ALGORITHMS FOR
CRYPTOCURRENCY FORECASTING**

I declare that the copyright holder of this thesis is International Islamic University Malaysia.

Copyright © 2022 Nor Azizah binti Hitam International Islamic University Malaysia. All rights reserved.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

4. Any material contained in or derived from this unpublished research may be used by others in their writing with due acknowledgement.
5. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purposes.
6. The IIUM library will have the right to make, store in a retrieved system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Nor Azizah binti Hitam

.....
Signature

.....
Date

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA
DECLARATION OF COPYRIGHT AND AFFIRMATION OF
FAIR USE OF UNPUBLISHED RESEARCH

SENTIMENT-BASED SUPPORT VECTOR MACHINE
OPTIMIZED BY METAHEURISTIC ALGORITHMS FOR
CRYPTOCURRENCY FORECASTING

I declare that the copyright holder of this thesis is Nor Azizah Binti Hitam

Copyright © 2022 Nor Azizah Binti Hitam. All rights reserved.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

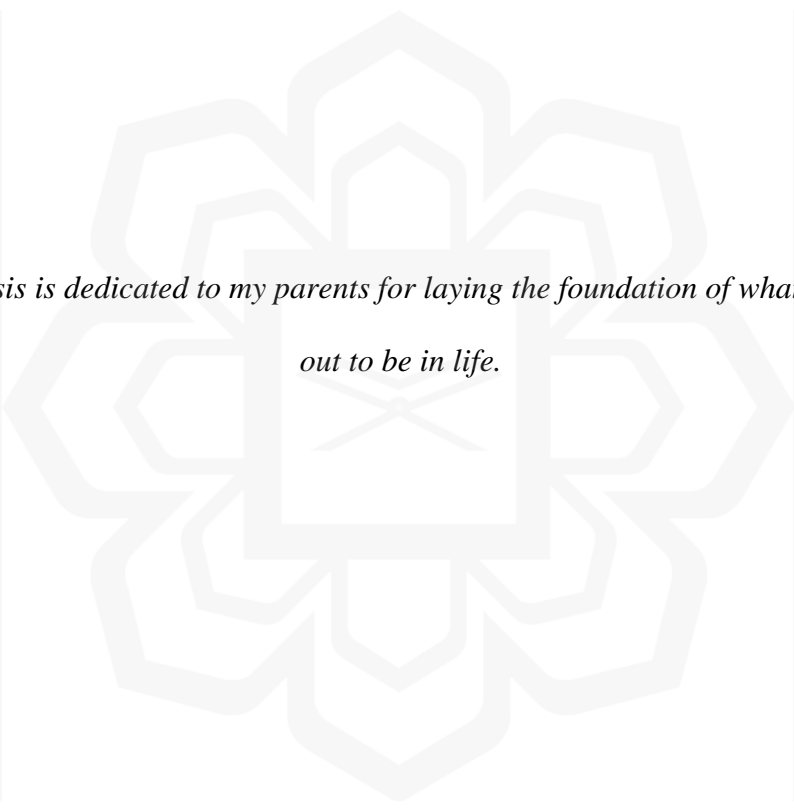
7. Any material contained in or derived from this unpublished research may be used by others in their writing with due acknowledgement.
8. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purposes.
9. The IIUM library will have the right to make, store in a retrieved system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Nor Azizah binti Hitam

.....
Signature

.....
Date



*This thesis is dedicated to my parents for laying the foundation of what I turned
out to be in life.*

ACKNOWLEDGEMENTS

All praises to Allah for the strengths and His blessings for the completion of this research and throughout the duration of my research.

I submit my heartiest gratitude to my respected supervisor I, Assoc. Prof. Dr Amelia Ritahani Ismail and supervisor II, Assoc. Prof. Dr. Noormaziah binti Abdul Aziz, for giving me the opportunity to and providing me invaluable guidance throughout this research. Not only that, I appreciate their emotional support, sincerity, understanding, patience, encouragement and for pushing me farther than I thought I could go. I am extremely grateful for what they have offered me. May Allah rewards them abundantly

My deepest gratitude goes to my beloved husband, Syed Kairil Musairi Syed Abdul Mutalib Jamalullail for his patience and companionship, and to my five wonderful children: Syed Ariq Darwish, Syed Aisy Fareeq, Syed Adel Harraz, Syed Aqil Jahran and the one and only girl, Sharifah Arwa Luthfia; for their love, understanding, prayers, and endless support throughout the process of working on this research. My parents, Hitam bin Muhammad, and Halimah binti Abdul Ghani, for their love, prayers, and passionate encouragement, made it possible for me to complete this research. Thank you to my siblings and in-laws for being there thick and thin..

Above ground, I am deeply indebted to my friends for all the moral support and encouragement through this PhD journey. Thank you for everything. This piece of work I dedicate to my parents, siblings, relatives, and friends for their emotional support and for making me believe in myself to keep going and complete this thesis. Thank you for reminding me the importance of working hard and working sincerely.

TABLE OF CONTENTS

ABSTRACT	ii
ABSTRACT IN ARABIC	iii
APPROVAL PAGE	iv
DECLARATION	v
COPYRIGHT	viii
DEDICATION	ix
ACKNOWLEDGEMENTS	x
TABLE OF CONTENTS	xi
LIST OF TABLES	xv
LIST OF FIGURES	xvii
LIST OF ALGORITHMS	xx
LIST OF ABBREVIATIONS	xxi
CHAPTER ONE: INTRODUCTION	1
1.1 OVERVIEW	1
1.2 PROBLEM BACKGROUND	2
1.3 PROBLEM STATEMENT	3
1.4 RESEARCH QUESTIONS	5
1.5 RESEARCH HYPOTHESES	6
1.6 RESEARCH OBJECTIVES	6
1.7 RESEARCH GOAL AND CONTRIBUTIONS.....	7
1.8 RESEARCH SCOPE AND LIMITATION	8
1.8.1 Scope	8
1.8.2 Limitations.....	8
1.9 THESIS STRUCTURE.....	8
CHAPTER TWO: LITERATURE REVIEW	10
6.4 CRYPTOCURRENCY	11
6.5 A REVIEW ON SENTIMENT ANALYSIS	14
6.6 TWITTER SENTIMENT ANALYSIS	17
2.3.1 Lexicon Based Approach	17
2.3.2 Machine Learning Approach.....	22
Grammatical Tagging.....	23
Naïve Bayes Classification.....	24
2.3.2.3 Support Vector Machine (SVM)	28
6.7 SENTIMENT CLASSIFICATION: TEXTBLOB	29
2.4.1 Regular Expression (REGEX/RE) Library	30
6.8 TIME SERIES FORECASTING ALGORITHM.....	31
2.5.1 Artificial Neural Network (ANN)	35
2.5.2 Deep Learning	39
2.5.3 Other Approaches	40
6.9 TECHNICAL INDICATORS.....	42
2.6.1 Commodity Channel Index (CCI)	43
6.10 SUPPORT VECTOR MACHINE (SVM).....	44
2.7.1 Kernel Function Selection	44

6.11	METAHEURISTIC ALGORITHMS.....	48
2.8.1	Genetic Algorithms (GA).....	48
2.8.2	Firefly Algorithm (FFA).....	50
2.8.3	Whale Optimization Algorithm (WOA).....	51
2.8.4	Particle Swarm Optimization Algorithm (PSO).....	53
2.8.5	Moth Flame Optimization Algorithm (MFO).....	62
	Begin.....	70
	End	70
6.12	RESEARCH GAP FINDING.....	70
6.13	CHAPTER SUMMARY.....	73

CHAPTER THREE: RESEARCH METHODOLOGY 74

3.1	PHASE ONE – LITERATURE REVIEW.....	77
3.2	PHASE TWO - DATA PREPARATION.....	77
3.2.1	Data Acquisition.....	77
3.2.2	Data Preprocessing.....	78
A)	Historical Data.....	78
B)	Sentiment Data.....	78
C)	Data Transformation.....	86
D)	Data Structuring.....	87
3.2.3	Benchmarks Data.....	87
3.2.4	Technical Indicator.....	87
3.3	PHASE THREE: EXPERIMENTAL DESIGN.....	88
3.3.1	Support Vector Machine (SVM).....	88
3.3.2	Metaheuristic Algorithms.....	89
3.4	PHASE FOUR: PERFORMANCE EVALUATION.....	89
3.5	PHASE FIVE: REPORT ON RESEARCH FINDINGS.....	90
3.6	RESEARCH MAP.....	90
3.7	CHAPTER SUMMARY.....	91

CHAPTER FOUR: EXPERIMENTAL DESIGN 92

4.1	FORECASTING MODEL.....	92
4.1.1	Data Construction.....	94
4.1.2	Data Normalization.....	105
4.1.3	Data Scaling.....	106
4.1.4	Proper Data Handling.....	106
4.2	THE SVM-KERNEL AND PARAMETERS.....	106
4.2.1	SVM Parameter Selection And Hyper Parameter Tuning.....	107
4.3	OTHER PARAMETERS.....	110
4.3.1	Population Size.....	111
4.3.2	Other Parameters.....	111
4.3.3	Stopping Criterion And The Number Of Iterations.....	111
4.4	SUPPORT VECTOR MACHINE WITH PARTICLES.....	112
4.5	SENTIMENT-BASED SUPPORT VECTOR MACHINE OPTIMIZED WITH METAHEURISTIC ALGORITHMS FOR CRYPTOCURRENCY FORECASTING.....	114
4.6	EXPERIMENTAL DATA.....	122

4.6.1	Bitcoin, BTC	123
4.6.2	Dash, DASH.....	123
4.6.3	Ethereum, ETH	123
4.6.4	Litecoin, LTC	124
4.6.5	Australian Dollar/US Dollar, AUDUSD.....	124
4.6.6	Euro/US Dollar, EURUSD.....	125
4.6.7	US Dollar/Canadian Dollar, USDCAD.....	125
4.6.8	British Pound/US Dollar, GBPUSD	125
4.6.9	New Zealand Dollar/US Dollar, NZDUSD	126
4.7	PERFORMANCE MEASURES AND EVALUATION METRICS.....	126
4.7.1	Performance Metrics	127
4.7.2	Statistical Correlation.....	127
4.7.3	Pearson’s Correlation Coefficient	127
4.7.4	Scatterplot	128
4.7.5	Error Accuracy	128
4.7.6	Paired T-Test.....	129
4.8	CHAPTER SUMMARY.....	130
CHAPTER FIVE: RESULT AND DISCUSSION		131
5.1	DESCRIPTIVE STATISTICS	132
5.2	VISUAL INSPECTIONS	132
5.3	NORMALITY TEST	137
	Normal	139
	Normal	140
5.4	CORRELATION ANALYSIS.....	141
5.4.1	Bivariate Correlation Of Twitter Sentiments And Cryptocurrency Market Price	141
5.4.2	Bivariate Correlation Of Twitter Sentiments And Forex Index Markets.....	149
5.5	EXPERIMENTAL RESULTS	158
5.6	PERFORMANCE METRIC RESULTS	160
5.6.1	Statistical Correlation	160
5.6.2	Error Accuracy	217
5.7	COMPARISON OF FORECASTING MODELS	222
5.7.1	RESULTS WITH AND WITHOUT SENTIMENT.....	222
	BTC, Bitcoin Price	222
	DASH, Dash Price.....	223
	ETH, Ethereum Price	223
	LTC, Litecoin Price.....	223
	AUDUSD Price	223
	EURUSD Price.....	224
	USDCAD Price	224
	GBPUSD Price	224
	NZDUSD Price.....	225
5.7.2	REAL CRYPTOCURRENCY DATA	225
	BTC Bitcoin Price	225
	DASH, Dash Price.....	227
	ETH Ethereum Price	229
	LTC Litecoin Price.....	231

5.7.3 REAL FOREX DATA	233
AUDUSD Price	233
EURUSD Price.....	235
USDCAD Price	237
GBPUSD Price	239
NZDUSD Price.....	241
5.8 TIME SERIES FORECASTING FOR BENCHMARK DATA:	
BITCOIN, BTC	244
5.8.1 Comparison Of The Proposed Model With Previous Literature	246
5.9 PAIRED T-TEST COMPARISON	247
5.9.1 Discussion On The Proposed Hybrid Sentiment-Based SVMPSOMFO Model.....	255
5.10 CHAPTER SUMMARY	259
CHAPTER SIX: CONCLUSIONS AND FUTURE WORKS	261
6.1 CONCLUSION.....	261
6.2 RESEARCH ACHIEVEMENTS	264
6.3 FUTURE WORKS.....	266
6.3 PUBLICATIONS.....	266
REFERENCES.....	269

LIST OF TABLES

Table 2.1 Moth Flame Optimization Terminology	65
Table 4.1 Description of the Selected Index	95
Table 4.2 Cryptocurrency data	101
Table 4.3 Forex data	101
Table 4.4 CCI sample Calculation	102
Table 4.5 Experimental Data	122
Table 5.1 Descriptive statistics for Cryptocurrency Price Index market	139
Table 5.2 Descriptive statistics for Forex market	140
Table 5.3 Bivariate Correlation of Twitter Sentiments and Cryptocurrency Market Price	141
Table 5.4 Bivariate Correlation of Twitter Sentiments and Forex Market Price	149
Table 5.5 Results of Performance Accuracy by Various Algorithms	160
Table 5.6 Summary Analysis for Statistical Correlation	216
Table 5.7 The Outcome of Indexes in Terms of MAE	218
Table 5.8 The outcome of indexes in terms of MSE	219
Table 5.9 The outcome of indexes in terms of RMSE	220
Table 5.10 The outcome of indexes in terms of MAPE	221
Table 5.11 Results with and without sentiment for Bitcoin	222
Table 5.12 Results with and without sentiment for Dash	223
Table 5.13 Results with and without sentiment for Ethereum	223
Table 0.14 Results with and without sentiment for Litecoin	223
Table 0.15 Results with and without sentiment for AUDUSD	223
Table 0.16 Results with and without sentiment for EURUSD	224
Table 5.17 Results with and without sentiment for USDCAD	224
Table 5.18 Results with and without sentiment for GBPUSD	224
Table 0.19 Results with and without sentiment for NZDUSD	225
Table 5.20 Comparative performances of all models for Bitcoin	225
Table 0.21 Comparative performances of all models for Dash	227
Table 5.22 Comparative performances of all models for ETH Ethereum	229
Table 5.23 Comparative performances of all models for LTC Litecoin	231
Table 5.25 Comparative performances of all models for EURUSD	235
Table 5.26 Comparative performances of all models for USDCAD	237

Table 5.27 Comparative performances of all models for GBPUSD	239
Table 5.28 Comparative performances of all models for NZDUSD	241
Table 5.29 Comparative performances of all models for benchmark data, BTC Bitcoin	244
Table 5.30 Comparison of performance between Hybrid sentiment-based SVMP SOMFO and Neuro Fuzzy Controller System (PATSO S).	247
Table 5.31 Paired T-Test Values of All Models for Bitcoin Data.	248
Table 5.32 Paired T-Test Values of All Models for Dash Data.	249
Table 5.33 Paired T-Test Values of All Models for Ethereum Data.	249
Table 5.34 Paired T-Test Values of All Models for Litecoin Data	250
Table 5.35 Paired T-Test Values of All Models for AUDUSD Data.	251
Table 5.36 Paired T-Test Values of All Models for EURUSD Data.	251
Table 5.37 Paired t-test values of all models for USDCAD data.	252
Table 5.38 Paired T-Test Values of All Models for GBPUSD Data.	252
Table 5.39 Paired T-Test Values of All Models for NZDUSD Data.	253
Table 6.1 Research Objectives Achievements	264

LIST OF FIGURES

Figure 2.1 Breakdown of the Chapter 2	10
Figure 2.2 Proposed Sentiment Analysis (SA) system architecture	19
Figure 2.3 Emotion Vector	21
Figure 2.4 Sentiment Calculation	21
Figure 2.5 Sentiment Analysis Modules Sentiment Analysis Modules	26
Figure 2.6 Framework of the web Crawler	29
Figure 2.7 The Support Vector Machine Optimal Hyperplane	45
Figure 2.8 Pseudocode of Genetic Algorithm (GA)	49
Figure 2.9 The Development of A Solution	50
Figure 2.10 Pseudocode for Firefly Algorithm (FFA)	51
Figure 2.11 WOA humpback whale bubble net feeding style	52
Figure 2.12 Velocity Vector	54
Figure 2.13 Flowchart of Particle Swarm Optimization (PSO) algorithm	59
Figure 2.15 The Spiral Flight Path of a Moth Around a Light Source	64
Figure 2.16 Flowchart of the Moth Flame Optimization algorithm (MFO)	66
Figure 2.17 Pseudocode of the Moth Flame Optimization	70
Figure 3.3 Python's Tweepy library	83
Figure 4.1 Design of Chapter 4	92
Figure 4.2 Graphs for cryptocurrency index and Forex index price movement over the training and testing period	100
Figure 4.3 Hyper Parameters Tuning	109
Figure 4.4 Flowchart of SVM with particles.	112
Figure 4.5 The flowchart of sentiment-based Support Vector Machine optimized with metaheuristic algorithms for cryptocurrency forecasting	114
Figure 5.1 Design of Chapter 5	132
Figure 5.2 Q-Q plots for cryptocurrency and Forex price index market datasets:	137
Figure 5.3 Cryptocurrency markets price and the sentiment Tweets. (a)	143

Figure 5.4 QQ plots (quantile-quantile plots) for data normality of Bitcoin market price and sentiment	145
Figure 5.5 QQ plots (quantile-quantile plots) for data normality of Dash market price and sentiment	146
Figure 5.6 QQ plots (quantile-quantile plots) for data normality of Ethereum market price and sentiment	147
Figure 5.7 QQ plots (quantile-quantile plots) for data normality of Litecoin market price and sentiment	148
Figure 5.8 Cryptocurrency markets price and the sentiment Tweets.	152
Figure 5.9 QQ plots (quantile-quantile plots) for data normality of AUDUSD market price and sentiment	153
Figure 5.10 QQ plots (quantile-quantile plots) for data normality of EURUSD market price and sentiment	154
Figure 5.11 QQ plots (quantile-quantile plots) for data normality of USDCAD market price and sentiment.	155
Figure 5.12 QQ plots (quantile-quantile plots) for data normality of GBPUSD market price and sentiment	156
Figure 5.13 QQ plots (quantile-quantile plots) for data normality of NZDUSD market price and sentiment.	157
Figure 5.14 Scatterplot for Bitcoin Datasets	167
Figure 5.15 Scatterplot for Dash Datasets	173
Figure 5.16 Scatterplot for Ethereum Datasets	179
Figure 5.17 Scatterplot for Litecoin Datasets	185
Figure 5.18 Scatterplot for AUDUSD Datasets	191
Figure 5.19 Scatterplot for EURUSD Datasets	197
Figure 5.20 Scatterplot for USDCAD Datasets	203
Figure 5.21 Scatterplot for GBPUSD Datasets	209
Figure 5.22 Scatterplot for NZDUSD Datasets	215
Figure 5.23 Graph of actual and predicted values for individual SVM models using Bitcoin data.	227
Figure 5.24 Graph of actual and predicted values for individual SVM models using Dash data.	229
Figure 5.25 Graph of actual and predicted values for individual SVM models using Ethereum data.	231

Figure 5.26 Graph of actual and predicted values for individual SVM models using Litecoin data.	233
Figure 5.27 Graph of actual and predicted values for Optimized SVM models of AUDUSD data.	235
Figure 5.28 Graph of actual and predicted values for Optimized SVM models of EURUSD data.	237
Figure 5.29 Graph of actual and predicted values for Optimized SVM algorithms of USDCAD data.	239
Figure 5.30 Graph of actual and predicted values for Optimized SVM models of GBPUSD data.	241
Figure 5.31 Graph of actual and predicted values for Optimized SVM models of NZDUSD data.	243
Figure 5.32 Graph of actual and predicted values for individual SVM models using real Bitcoin data.	245
Figure 5.33 Graph of actual and predicted values for best Optimized SVM models, benchmark and proposed model using Bitcoin data.	245

LIST OF ALGORITHMS

Support Vector Machine (SVM)

Particle Swarm Optimization (PSO)

Moth Flame Optimization (MFO)

Support Vector Machine Particle Swarm Optimization Moth Flame
Optimization (SVMPSOMFO)



LIST OF ABBREVIATIONS

ALO	Ant Lion Optimization
FFA	Firefly Algorithm
Forex	Foreign Exchange
GA	Genetic Algorithm
GOA	Grasshopper Optimization Algorithm
GWO	Grey Wolf Optimization
HS	Harmony Search
MFO	Moth Flame Optimization
NLP	Natural Language Processing
PSO	Particle Swarm Optimization
SVM	Support Vector Machine
SVMALO	Support Vector Machine with Ant Lion Optimization
SVMFFA	Support Vector Machine with Firefly Algorithm
SVMGA	Support Vector Machine with Genetic Algorithm
SVMGOA	Support Vector Machine with Grasshopper Optimization
Algorithm	
SVMGWO	Support Vector Machine with Grey Wolf Optimization
SVMHS	Support Vector Machine with Harmony Search
SVMMFO	Support Vector Machine with Moth Flame Optimization
SVMPSO	Support Vector Machine with Particle Swarm Optimization
SVMPSOMFO	Support Vector Machine with Particle Swarm Optimization and Moth Flame Optimization
SVMWOA	Support Vector Machine with Whale Optimization Algorithm
WOA	Whale Optimization Algorithm

CHAPTER ONE

INTRODUCTION

1.1 OVERVIEW

Forecasting plays a vital role in decision making process and planning in any organizations. It can be described as a prediction of a future events based on the historical data. This includes time series forecasting. As economic trends and market trends are inseparable, economists considered bond and stock markets anticipate economic trends. Market analysts look at the market trends more than the economic trends. Fundamental analysts must be specialized on fewer market sectors and need to understand the nature of industries and thus demand more work and concentration including the sentiment of investors. Investors sentiment do influence by the textual analysis that are available on most of the social media such as mood or sentiment information about stock market price (Bollen et al., 2011; Brendan O'Connor et al., 2010a), public opinions and sales forecast (Brendan O'Connor et al., 2010b).

Profit of all investors and traders are depending on the level of predictability, predicting future market will help them in making corrective measures in their forecasting. A huge growth in computing power have revolutionized virtually every angle of modern live, and financial market are no exception. Financial market is known as one of the major contributors to the capital budget (Baker & Wurgler, 2013) and many researchers show that its development give direct impact to the financial growth (Ake & Ognaligui, 2010). Various models have been introduced by previous researchers in conducting time series forecasting including the development of the Artificial Intelligence (AI) models that leads to an efficient stock market and give more confidence to investors to allocate more capital, finance and accommodate the products exchange (Ake & Ognaligui, 2010). This research aims to design novel algorithms for cryptocurrency future price forecasting that is capable to deal with

huge datasets. To achieve this aim, the proposed algorithm is inspired from sentiment-based Support Vector Machines (SVM) and optimized by particle swarm and moth flame as a methodology in developing the model. This thesis review most of the machine learning algorithms that are broadly used by the researchers in forecasting. There are many types of machine learning algorithms including artificial neural networks, support vector machines, deep learning and other approaches. This chapter will explain about some motivation of this thesis with a brief explanation on technical and fundamental analysis, financial technology along with the background of the study. Following this, the structure and content of the thesis is developed, stating the objective, research questions, significance and scope including limitation while working towards the completion.

1.2 PROBLEM BACKGROUND

Internet users may use social media to pursue common interests and to connect with others, sharing opinions, allow communication as well as finding the existing acquaintances (Nisar & Yeung, 2018; Syed Zulkarnain & Hitam, 2014; Wikipedia Contributors, 2017). To date, many studies being done in cryptocurrency market forecasting due to its popularity (Indera et al., 2017; Lee Kuo Chuen, 2015; Peng et al., 2018; Roche & McNally, 2016; Uzeki, 2018). A study by Parikh et al. (2021) discovers that the sentiment does play an important role in the cryptocurrency market as the high engagement of the investors to the social media do influence the transaction of the cryptocurrency market during the particular period.

However, apart from the engagement of the investors to the social media, there are other factors that can be considered in forecasting the cryptocurrency prices. First is the development of an efficient model (Kumar Dash et al., 2021). The development of an efficient model to forecast future cryptocurrency prices is still difficult due to its complexity in development and operations (Indera et al., 2017; Kara et al., 2011; Kumar Dash et al., 2021, Roche & McNally, 2016). The issue with the development of the model are related to the hyperparameter

optimization and the feature selection as mentioned by (Azmira et al., 2017; Buslim et al., 2021; Huerta et al., 2013; Kara et al., 2011). Huerta et al., (2013), has designed a model and simulate series of experiments to test the proposed model on the selected classifiers. In accordance with the proposed model, smallest observation set of data contribute to better performance as the trainings in SVM relies on the available data, but it draws an attention on the number of trainings needed by the classifiers to is left unknowns. This is also explored by Kara et al. (2011) where the proposed model is to forecast a movement of stock price by applying machine learning algorithms which are ANN and SVM in Istanbul Stock Exchange. The model has outperformed other models; however, the shortcomings are the model disregards the hyperparameter optimization and feature selection. This is followed by another research in financial forecasting presents that even though SVM does reduce the local minima issues and over-fitting, high complexity in its programming that pays more attention to inequality compared to the equality constraints (Azmira et al., 2017) becomes another deficiency to the SVM. (J. Wang, 2017) finds that the constraints of SVM in terms of parameter optimization and feature selection must be examined while developing a forecasting model, as revealed in the research of (Buslim et al., 2021). A study of Uras and Ortu (2021) discovers that additional of technical indicators to the classic microeconomic variables leads to an effective improvement in the prediction of cryptocurrency performances. However, no one knows which indicators can perform best when implemented in the machine learning algorithm.

1.3 PROBLEM STATEMENT

The Support Vector Machine (SVM) demonstrates an increase in forecasting the future price for it has the potential to lessen the local minima problem and over-fitting issue, nevertheless, it disregards the hyperparameter optimization and feature selection (Huerta et al., 2013), demand high complexity in its programming which demands more attention to inequality compared to the equality constraints (Azmira et al., 2017). A recent study of (Uras & Ortu, 2021) proved that the SVM is still a relevant machine learning algorithm for