





PROJECT REPORT

Stock Prediction Market Using C4.5 Algorithm

Dandhi Widiadamarrama

09.02.0005

2013

	PERPUSTAKAAN
NO. INV : 241 / S / IK / C . 1	
TGL : 06 10 2014	
PARAF : 	

FACULTY OF COMPUTER SCIENCE
SOEGIJAPRANATA CATHOLIC UNIVERSITY

Jl. Pawiyatan Luhur IV/1, Bendan Duwur, SEMARANG 50234

Telp. 024-8441555 (hunting) Web: <http://www.unika.ac.id>

<http://ikomunika.web.id/>

APPROVAL AND RATIFICATION PAGE

PROJECT REPORT

Stock Prediction Using C4.5 Algorithm

This project report has been approved and ratified by the Dean of Faculty of Computer Science and Supervisor on January, 21th 2014

With approval,

Examiners,



Suyanto Edward Antonius, Jr., M.Sc
NPP : 058.1.1992.116

Supervisor,



Hironimus Leong, S.Kom., M.Kom
NPP : 058.1.2007.273

Examiners,



Shinta Estri Wahyuningrum, S.Si, M.Cs
NPP : 058.1.2007.272

Examiners,



Rosita Herawati, ST., MIT
058.1.2004.263



Dean of Faculty of Computer Science,



Hironimus Leong, S.Kom., M.Kom
NPP : 058.1.2007.273

STATEMENT OF ORIGINALITY

I, the undersigned:

Name : Dandhi Widiadamarrama

ID : 09.02.0005

Certify that this project was made by myself and not copy or plagiarize from other people, except that in writing expressed to the other article. If it is proven that this project was plagiarizes or copy the other, I am ready to accept a sanction.

Semarang, January 21th 2014



Dandhi Widiadamarrama

09.02.0005

ABSTRACT

Stock market prediction with data mining techniques is one of the most important issues to be investigated. In this market, stock currency are sold and bought based on the exchange rate. It is important to participants to accurately predict future values and trends. In this project,writer's application has been designed and predict stock change values and trends using data mining techniques and algorithms.

One of the methods that used in this application is called Decision Tree and C 4.5 as algorithm. Before the analyzing, the historical stock data must through across the '*data cleaning*' and '*data construct*' process. To demonstrate that this application is working correctly, writer used the prediction system to analyze and stock data to predict the price promptly at the specific time. The application can effectively help trader deal with day trading.

The goal of this research is to gather stock market prediction based on decision tree that build up from the application. The prediction will be presented in the form of decision tree from the attribute of stock data as an decision system making. Despite of many benefits that we can get from this application, there are limitations that should be investigated,such as relevance of the result and the best "*topology*" for certain problems.

keywords : stock prediction ,decision tree, c4.5 algorithm

FOREWORD

Praise to the God for all of his blessings that keep me spirited and inspired all days to finish my thesis entitled Stock Prediction Market Using C4.5 Classification Decision Tree Algorithm . Also, without the role of these people, I would not be able to finish my thesis; I would like to thank to:

1. My parents and my family in for all of their endless prayers and supports for me
2. My supervisor: Mr. Hironimus Leong, S.Kom., M.Kom for his endless supports, guides and suggestions for my thesis
3. All of my friends in Faculty of Computer Science, as well as any other friends of mine outside Faculty of Computer Science who keep motivate me.

I would like to apologize if there are some drawbacks in my projects. Hopefully this project can be useful for other people.

Semarang, January 21, 2014

Dandhi Widiadamarrama

09.02.0005

Table of contents

APPROVAL AND RATIFICATION PAGE.....	i
STATEMENT OF ORIGINALITY.....	ii
ABSTRACT.....	iii
FOREWORD	iv
TABLE OF CONTENTS.....	v
TABLE OF FIGURES.....	vii
TABLE OF TABLES	viii
CHAPTER I	1
CHAPTER II.....	3
2.1 Stock.....	3
2.1.1 Historical Prices.....	4
2.1.2 Yahoo Finance API.....	5
2.2 Data Mining.....	5
2.2.1 Introduction to Data Mining.....	5
2.2.2 Data Mining Purpose	5
2.2.3 Data Mining Process.....	6
2.2.4 C4.5 Algorithm.....	8
2.2.5 Decision Tree	12
CHAPTER III.....	13
CHAPTER IV	14
4.1 Analysis	14

4.1.1 Use Case Diagram	14
4.2 Design.....	15
4.2.1 Flow Chart Diagram	15
4.2.2 Class Diagram	16
CHAPTER V.....	16
5.1 Implementation	16
5.1.1 Download Historical Stock Data	16
5.1.2 Data Cleaning Process	19
5.1.3 Mining C4.5	22
5.2. Testing	28
5.2.1 Testing Table.....	28
5.2.2 Testing Screenshot.....	30
CHAPTER VI	37
6.1 Conclusion.....	37
6.2 Further Research.....	37
REFERENCES.....	38

Table of figures

Figure 2.3 Gain Formula.....	9
Figure 5.1.3Index.php screenshot.....	17
Figure 5.1.4 Stock Downloader Code.....	18
Figure 5.1.5 Data Cleaning Process.....	20
Figure 5.1.6 createView.php Code	21
Figure 5.1.7 Cleaning Stored Procedure	21
Figure 5.1.8 Mining C4.5.....	22
Figure 5.1.9 Calculate entropy total.....	23
Figure 5.1. 10 Calculate entropy attribute.....	24
Figure 5.1.12 Calculate Gain Ratio	25
Figure 5.1.13 Insert tree	26
Figure 5.1.15 Recursive Loop	26
Figure 16 Replace Null	27
Figure 5.1.17 Populate Attribute	27
Figure 5.1.18 Create Table.....	28

TABLE OF TABLES

Table 3.1 Project management timeline.....	14
Table 5.1 Testing table.....	29