

**SENTIMENT ANALYSIS USING AUTOMATIC CLASSIFICATION ON  
ONLINE MEDIA ARTICLE**

By

Feizal Badri Asmoro  
12109010

A thesis submitted to the Faculty of  
ENGINEERING AND INFORMATION TECHNOLOGY

In Partial Fulfillment of the Requirements  
for

BACHELOR'S DEGREE

in

INFORMATION TECHNOLOGY

SWISS GERMAN UNIVERSITY



SWISS GERMAN UNIVERSITY  
EduTown BSD City  
Tangerang 15339  
Indonesia

Revision after Thesis Defense: 22 July 2013

## STATEMENT BY THE AUTHOR

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at any educational institution, except where due acknowledgement is made in the thesis.

Feizal Badri Asmoro

Student

Date

Approved by:

James Purnama

Thesis Advisor

Date

Alva Erwin, M. Sc.

Thesis Co-Advisor

Date

Dr. Ir. Gembong Baskoro, M. Sc

Dean

Date

## ABSTRACT

### SENTIMENT ANALYSIS USING AUTOMATIC CLASSIFICATION ON ONLINE MEDIA STREAMING

By

Feizal Badri Asmoro  
James Purnama. M.Sc, Advisor  
Alva Erwin, M.Sc, Co-Advisor

SWISS GERMAN UNIVERISTY

The increasing demand for an automatic sentiment measurement is becoming critical in this era. Therefore, this topic is becoming one of the most encountered reseraches in data mining and NLP. Sentiment analysis is a process of classifying articles as a positive or negative. This kind of approach is lacking of multiple assessor's valuation. In order to do so, a more humanly semantic classification is needed. There are various researches on developing an algorithm for sentiment analysis by using classification, such as research by Turney (2010) and Pang (2008). The main purpose of this research is to design and implement a system that can analyze sentiment of text in Bahasa Indonesia. The article will be categorized into one of two sentiments: positive or negative. The text is an online digital media streaming article. Sentiment analysis will use automatic classification to analyze the sentiment.

*Keywords: sentiment, natural language processing, wordnet, automatic classification.*



**SWISS GERMAN UNIVERSITY**

## DEDICATION

I dedicate this works to my parents for supporting me on this research, friends who always helped me with their valuable knowledge and all researchers who are working in NLP.



## ACKNOWLEDGEMENTS

I owe a great many thanks to a great many people who helped and supported me during the working on this research.

My deepest thanks to James Purnama, M.Sc and Alva Erwin, M.Sc for guiding and correcting the works with full attention. They have taken pain to go through the research and make necessary correction.

I would like also to thank my Institution and my faculty members without whom this project would have been a distant reality. I also extend my heartfelt thanks to my family and well-wishers.



SWISS GERMAN UNIVERSITY

## TABLE OF CONTENTS

STATEMENT BY THE AUTHOR.....	2
ABSTRACT .....	3
DEDICATION .....	5
ACKNOWLEDGEMENTS.....	6
TABLE OF CONTENTS .....	7
LIST OF FIGURES.....	9
LIST OF TABLES.....	10
<b>CHAPTER 1 - INTRODUCTION .....</b>	<b>11</b>
1.1 Background .....	11
1.2 Objectives.....	12
1.3 Research Limitation .....	12
1.4. Research Problem .....	12
1.5 Significance of Study.....	13
1.6 Research Questions and Hypothesis .....	13
1.7 Methodology.....	14
<b>CHAPTER 2 - LITERATURE REVIEW.....</b>	<b>16</b>
2.1 Sentiment Analysis.....	16
2.2 Natural Language Processing.....	17
2.3 WordNet.....	22
2.4 K-nearest Neighbors Algorithm for Data Classification.....	23
2.5 Selenium Framework .....	26
2.5.1 Selenium IDE.....	26
2.5.2 Selenium Client API .....	26
2.5.3 Selenium Remote Control.....	27
2.5.4 Selenium WebDriver.....	27
2.5.5 Selenium Grid.....	27
2.5.6 Web Element Locator and Web Element Actions.....	27
<b>CHAPTER 3 - METHODOLOGY .....</b>	<b>29</b>
3.1 Text Crawler Framework .....	29
3.1.1. Crawling <a href="http://kamusbesar.com">http://kamusbesar.com</a> .....	29
3.1.2. Crawling <a href="http://sinonimkata.com">http://sinonimkata.com</a> .....	31
3.1.3. Crawling <a href="http://translate.google.com">http://translate.google.com</a> .....	32
3.2 Text Parser .....	34
3.3 Wu & Palmer Semantic Measurement.....	38
3.4 Similarity Matrix .....	41
<b>CHAPTER 4 – RESULT AND DISCUSSION.....</b>	<b>44</b>
4.1. Overview.....	44
4.2. Wu & Palmer .....	44

---

<b>4.3 Similarity Matrix</b> .....	<b>46</b>
<b>4.4 Automatic classification using k-Nearest Neighbor</b> .....	<b>48</b>
<b>4.5 Evaluation</b> .....	<b>53</b>
<b>CHAPTER 5 – CONCLUSION AND RECOMMENDATION</b> .....	<b>55</b>
<b>5.1 Conclusion</b> .....	<b>55</b>
<b>5.2 Recommendation</b> .....	<b>56</b>
<b>GLOSSARY</b> .....	<b>57</b>
<b>REFERENCES</b> .....	<b>58</b>
<b>APPENDIX</b> .....	<b>60</b>
<b>CURRICULUM VITAE</b> .....	<b>64</b>

