

**DEVELOPMENT OF SENTIMENT ANALYSIS OF INSTAGRAM
COMMENTS FOR SOCIAL MEDIA INFLUENCER ONLINE
MARKETPLACE**

By

Davega Putra Prasisto
11402011



SWISS GERMAN UNIVERSITY

The Prominence Tower

Jalan Jalur Sutera Barat No. 15, Alam Sutera

Tangerang, Banten 15143 - Indonesia

Revision after the Thesis Defense on July 19th 2019

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.

Davega Putra Prasisto

Student

Date

Approved by:

James Purnama M.Sc

Thesis Advisor

Date

Dipl.-inf. Kho I Eng., M.Sc.

Thesis Co-Advisor

Date

Dr. Maulahikmah Galinium, S.Kom.

Dean

Date

Davega Putra Prasisto

ABSTRACT

DEVELOPMENT OF SENTIMENT ANALYSIS OF INSTAGRAM COMMENTS FOR SOCIAL MEDIA INFLUENCER ONLINE MARKETPLACE

By

Davega Putra Prasisto

James Purnama M.Sc, Advisor

Dipl.-inf. Kho I Eng.,M.Sc., Co-Advisor

SWISS GERMAN UNIVERSITY

Influencers is increasingly being utilized in word of mouth social media marketing and various online platforms has appeared that serves as a marketplace that connect clients with influencers and offering various services, including assessing the general perception of the marketing campaign. Such assessment is made by analyzing the comments affiliated with the social media posts by influencers of the marketing campaign, which could be difficult to perform manually due to the potentially large amount of comments. As such, this research proposes the development of a lexicon-based sentiment analysis system for social media influencer online marketplace. The sentiment analysis system would be able to differentiate Instagram comments that are irrelevant to the Instagram post and would be able to classify the sentiment polarity. The sentiment analysis system is developed based on waterfall software development life cycle model and subjectivity measurement is adopted as an approach to measure its performance as indicated by four different metrics: Accuracy, Precision, Recall, and F1-measure. The result show that the proposed sentiment analysis system could reliably classify Instagram comments that is relevant to its affiliated Instagram post and express positive sentiment, but improvements need to be made on the classification on other categories.

Keywords: Sentiment Analysis, Influencer, Waterfall Methodology, SentiStrength, Lexicon-based Approach.



DEDICATION

I dedicate to the advisors and lecturers of Swiss German University whose help has been essential the completion of this work, this work to my dear parents who have always loved and supported me, to my friends, to the future and well-being of my country: Indonesia, and finally to Allah Subhanahu wa ta'ala.



ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to the advisors and lecturers of Swiss German University since without their help and guidance this work could not be completed. Special thanks to Sir James Purnama as my advisor in this work, whose encouragement, optimism, and patience has been vital for the completion of this research. I also would like to thank Sir Kho I Eng as my co-advisor for the advices that kindly provided.

Second, I would also like to thank my parents, especially my dear father Dede Soeprapto. He has always loved and supported me even when I am at my lowest point, come hell or high water. and want nothing but the best for me. As such, I am eternally grateful for him and hope that this work and its results would be able to please him and make him happy.

Last but not least, I extend my gratitude to my friends both former college classmates who had graduated ahead of me and those from my high school days who has provided me with entertainment, encouragement, and advises that greatly helped and encouraged me, even as I rarely met face-to-face with them during this work. Special thanks to Katon Pijar Dwisetio, my best friend from the senior high school whose encouragement, company, and good humor helps me keep my head up in this research.

TABLE OF CONTENTS

	Page
ABSTRACT.....	3
DEDICATION	5
ACKNOWLEDGEMENTS	6
TABLE OF CONTENTS	7
LIST OF FIGURES	9
LIST OF TABLES	10
CHAPTER 1 - INTRODUCTION.....	11
1.1 Background	11
1.2 Research Problems.....	14
1.3 Research Objective	14
1.4. Research Questions	14
1.5. Research Hypothesis.....	15
1.6 Research Scope	15
1.7 Research Limitation	15
CHAPTER 2 - LITERATURE REVIEW	17
2.1 Social Media Marketing.....	17
2.2 Influencers and Followers.....	17
2.3 Natural Language Processing.....	19
2.4 Sentiment Analysis	24
2.4.1 Sentiment Analysis Approaches.....	28
2.4.2 Levels of Sentiment Analysis	34
2.4.3 Challenges in Sentiment Analysis	35
2.5 Subjectivity Measurement and Performance Evaluation.....	38
2.5.1 Confusion Matrix.....	38
2.5.2 Accuracy	39
2.5.3 Precision and Recall	40
2.5.4 F1-measure	41
2.6 SentiStrength.....	41
2.7 Related Works on Sentiment Analysis	45
2.7.1 A Sentiment Analysis on Miley Cyrus' Instagram Accounts.....	45

2.7.2 A Sentiment Analysis Tool for Determining the Promotional Success of Fashion Images on Instagram	46
2.7.3 Extractive Sentiment Summarization Using Hybrid TF-IDF and Cosine Similarity	47
2.7.4 Predicting Brexit: Classifying Agreement is Better than Sentiment and Pollsters	48
2.7.5 Sentiment Analysis Using Automatic Classification on Online Media Article	49
2.7.6 Sentiment Analysis of Indonesian Low Cost Green Cars with Twitter Data	50
2.7.7 Sentiment Analysis of Indonesian Mobile Operator with Twitter Data	51
2.8 Instagram.....	52
2.9 Software Development Life Cycle.....	52
CHAPTER 3 – RESEARCH METHODOLOGY	58
3.1 System Analysis, Design, Development, and Testing Methodology.....	58
3.2 Sentiment Analysis Development Methodology	60
3.2.1 Sentiment Analysis System Requirement Analysis	60
3.2.2 Sentiment Analysis Design Creation.....	61
3.2.3 Sentiment Analysis System Development.....	63
3.2.4 Sentiment Analysis Experimentation	64
CHAPTER 4 – SYSTEM ANALYSIS AND DESIGN.....	65
4.1 System Requirements.....	65
4.2 Proposed System	67
4.2.1 SentiStrength as a Model for Lexicon-based Sentiment Analysis System	68
4.2.2. Architecture Diagram	70
4.2.3. Sentiment Analysis Process.....	71
4.2.4 Sentiment Lexicon.....	81
CHAPTER 5 – SYSTEM TESTING AND PERFORMANCE EVALUATION.....	85
5.1 Generating Test Data	85
5.2 Data Analysis	86
CHAPTER 6 – RESULTS AND DISCUSSIONS.....	88
6.1 Subjectivity Validation	88
6.2 Performance Measurement – Relevancy Classification.....	90
6.3 Performance Measurement – Sentiment Polarity Classification	92
CHAPTER 7 – CONCLUSIONS AND RECCOMENDATIONS	96
7.1 Conclusions.....	96
7.2 Limitation, Recommendation, and Future Works.....	97
REFERENCES	99
CURRICULUM VITAE.....	104