PROJECT REPORT

ANALYZING AND PREDICTING SENTIMENT RESULT FOR AMAZON PRODUCTS USING TEXTBLOB AND SUPPORT VECTOR MACHINE

TAN, DITYA YOSAPUTRA SISWOKO
18.K1.0018

Faculty of Computer Science
Soegijapranata Catholic University
2022

HALAMAN PENGESAHAN

Judul Tugas Akhir:	:	Analyzing and predicting sentiment result for Amazon products using TextBlob and Support Vector Machine
Diajukan oleh	:	Tan, Ditya Yosaputra Siswoko
NIM	:	18.K1.0018
Tanggal disetujui	:	15 Juli 2022
Telah setujui oleh		
Pembimbing	:	Yulianto Tejo Putranto S.T., M.T.
Penguji 1	:	Yonathan Purbo Santosa S.Kom., M.Sc
Penguji 2	:	Yulianto Tejo Putranto S.T., M.T.
Penguji 3	:	Hironimus Leong S.Kom., M.Kom.
Penguji 4	:	R. Setiawan Aji Nugroho S.T., MCompIT., Ph.D
Penguji 5	:	Rosita Herawati S.T., M.I.T.
Penguji 6	:	Y.b. Dwi Setianto S.T., M.Cs.
Ketua Program Studi	:	Rosita Herawati S.T., M.I.T.
Dekan	:	Dr. Bernardinus Harnadi S.T., M.T.

Halaman ini merupakan halaman yang sah dan dapat diverifikasi melalui alamat di bawah ini.

sintak.unika.ac.id/skripsi/verifikasi/?id=18.K1.0018

HALAMAN PERNYATAAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan dibawah ini:

Nama : Tan, Ditya Yosaputra Siswoko

Program Studi : Teknik Informatika

Fakultas : Ilmu Komputer

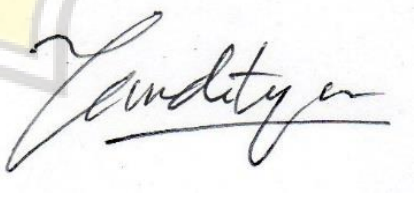
Jenis Karya : Skripsi

Menyetujui untuk memberikan kepada Universitas Katolik Soegijapranata Semarang Hak Bebas Royalti Nonekslusif atas karya ilmiah yang berjudul **"**Analyzing and predicting sentiment result for Amazon products using TextBlob and Support Vector Machine**"** beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Nonekslusif ini Universitas Katolik Soegijapranata berhak menyimpan, mengalihkan media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir ini selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Semarang, 23 Juli 2022

Yang menyatakan



Tan, Ditya Yosaputra Siswoko

DECLARATION OF AUTHORSHIP

I, the undersigned:

Name : Tan, Ditya Yosaputra Siswoko

ID : 18.K1.0018

declare that this work, titled “ANALYZING AND PREDICTING SENTIMENT RESULT FOR AMAZON PRODUCTS USING TEXTBLOB AND SUPPORT VECTOR MACHINE", and the work presented in it is my own. I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at Soegijapranata Catholic University
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
3. Where I have consulted the published work of others, this is always clearly attributed.
4. Where I have quoted from the work of others, the source is always given.
5. Except for such quotations, this work is entirely my own work.
6. I have acknowledged all main sources of help.
7. Where the work is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Semarang, July 23, 2022

Tan, Ditya Yosaputra Siswoko

18.K1.0018

ACKNOWLEDGMENT

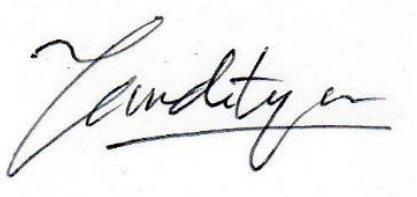
Praise to God the Almighty for His gift and mercy that I have finished tesis titled "ANALYZING AND PREDICTING SENTIMENT RESULT FOR AMAZON PRODUCTS USING TEXTBLOB AND SUPPORT VECTOR MACHINE" as the final requirement in accomplishing undergraduate at Informatics Engineering Departement Computer Science Faculty on Soegijapranata Catholic University Semarang.

For all the challenges and the difficulties I faced during on accomplishing this tesis, with any supports and guidance from any sides, so I should express my gratitude to :

1. My supervisor, Mr. Yulianto Tejo Putranto for supports, guides and directs me to finishing my project patiently.
2. All of lecturer from Informatics Engineering Departement who have teach, educate, shape, and motivate me into I am now to give me a way to the future on my study.
3. All of my family and relatives who supports me on undergoing study period with both materially and morally so that I can study well without any problems for giving me the best on my future.

Though of my limitations and mistake on this creation and research. Hopefully the constructive critics and suggestions could perfecting this tesis and encourage us to able to perfect and better research in the future.

Semarang, July 23, 2022



Tan, Ditya Yosaputra Siswoko

ABSTRACT

Nowadays, numerous of reviews from buyers spreaded widely on online shopping platform, especially Amazon which it give an important information for knowing how good the condition and service of this item from various available categories. The bulk of reviews from buyers complicate to indentify all of sentiments of many comments, so that program was made to identify and analyze all of sentiments of many comments using sentiment analysis method.

This program identifies sentiment from each of reviews based of average valuation from automatic polarity of reviews using TextBlob library from Python programming language and the overall of each comments which changed into certain numbers such as -1, -0,5, 0, 0,5, 1 for 1, 2, 3, 4, 5 overalls. If there are a comment which have helpful votes, so there are additional changed overall values to count the average of this sentiment valuation. The average valuation of sentiment will determine whether sentiment are positive or negative. Then the sentiment result of reviews are processed into Support Vector Machine (SVM) algorithm which later measuring accuracy level and predict other reviews from another prediction dataset to prove exactness of it.

From the project result that performed, the Polynomial Kernel from SVM algorithm is the highest accuracy of 92,85% on prediction result in Without Stemming experiment and for highest examination of sentiment identification prediction accuracy is Radial Basis Function (RBF) from SVM algorithm with 92,64% which 667 of 720 from another dataset are correctly predicted in Without Undersampling experiment.

Keyword: Sentiment Analysis, TextBlob library, Support Vector Machine, Python programming language

TABLE OF CONTENTS

LIST OF FIGURE

LIST OF TABLE