

KLASIFIKASI KEPRIBADIAN PENGGUNA FACEBOOK MENGUNAKAN SUPPORT VECTOR MACHINE

by Muhammad Yasa

FILE	JURNAL_YASA.DOCX (667.28K)	WORD COUNT	2591
TIME SUBMITTED	08-JAN-2020 09:46AM (UTC+0700)	CHARACTER COUNT	15925
SUBMISSION ID	1239930243		

KLASIFIKASI KEPERIBADIAN PENGGUNA FACEBOOK MENGUNAKAN SUPPORT VECTOR MACHINE

Muhammad Yasa¹ Elsen Ronando²

¹²
Fakultas Teknik Informatika Universitas 17 Agustus 1945 Surabaya,
Jl. Semolowaru no. 45 Surabaya, 60118, Telp : (031) - 5931800, Fax: (031) – 5927817,
E-Mail: muhammad.yasa92@gmail.com, elsen.ronando@untag-sby.ac.id

Abstract

³
Facebook is a social media that is used to socialize, from the status of Facebook users can be used to obtain information about the user's personality. ⁶ this study the personality taken consisted of social words, positive emotions and negative emotions. Linguistic Inquiry and Word Count (LIWC) is a program that can analyze and convert text into a value so that it can describe some of the values of one's personality.

This research use for develop the application 'personality classification of Facebook social media users', and to apply the SVM method in the classification process. This research focused on its application in the industrial world to select or selection process on prospective employees with status data from Facebook social media.

In this research, taking facebook status data as many as 300 status at random and as a comparison using data from the status of Twitter users as many as 2067 status. In the testing process using the Support Vector Machine (SVM) method using the normalization scheme [0.1] and [-1.1], the scheme [0.1] obtained 100% accuracy in the Facebook user status test and 70.29% for the Twitter user status test, and at scheme [-1.1] obtained 100% accuracy value for facebook user status and 99.3% accuracy value for twitter user status. From the test results found 100% value on the facebook status test and the test results on the status of Twitter users found a value of 70.29% for the scheme [0.1] and an accuracy value of 99.3% in the scheme [-1.1].

Keywords: Facebook, Support Vector Machine, Linguistic Inquiry and Word Count (LIWC), Classification

Abstrak

Facebook merupakan sosial media yang dipergunakan untuk bersosialisasi, dari status pengguna facebook dapat dipergunakan untuk memperoleh informasi kepribadian penggunanya. Pada penelitian ini kepribadian yang diambil terdiri dari social word, positive emotions dan negative emotions. Linguistic Inquiry and Word Count (LIWC) adalah program yang dapat menganalisa dan merubah teks menjadi sebuah nilai sehingga dapat menggambarkan beberapa nilai-nilai kepribadian seseorang.

Penelitian ini bertujuan untuk mengembangkan aplikasi 'klasifikasi kepribadian pengguna sosial media facebook', serta menerapkan metode SVM pada proses klasifikasinya. Penelitian ini berfokus pada penerapannya di dunia industri untuk menyeleksi atau proses seleksi pada calon pegawai dengan data status dari sosial media facebook.

Pada penelitian ini mengambil data status pengguna facebook sebanyak 300 status secara acak dan sebagai pembandingan menggunakan data dari status pengguna twitter sebanyak 2067 status. Pada proses pengujian menggunakan metode Support Vector Machine (SVM) memakai skema normalisasi [0.1] dan [-1.1], pada skema [0.1] didapatkan nilai akurasi 100% pada uji status pengguna facebook dan nilai 70.29% untuk uji status pengguna twitter, dan pada skema [-1.1] didapatkan nilai akurasi 100% untuk status pengguna facebook dan nilai akurasi 99.3% pada status pengguna twitter. Dari hasil pengujian yang dilakukan didapati nilai 100% pada pengujian status facebook dan hasil pengujian pada status pengguna twitter didapati nilai 70.29% untuk skema [0.1] dan nilai akurasi 99.3% pada skema [-1.1].

Kata Kunci: Facebook, Support Vector Machine, Linguistic Inquiry and Word Count (LIWC), Klasifikasi

1. PENDAHULUAN

Salah satu cara yang dipergunakan di dunia psikologi untuk mengetahui watak perilaku dari seseorang adalah dengan melihat kepribadiannya, terutama pada ruang lingkup pekerjaan agar didapati hubungan yang sesuai. Kepribadian atau watak dari seseorang dapat diketahui dari banyak aspek seperti kepribadian introvert, extrovert dan sebagainya. Pengukuran jenis kepribadian seseorang dapat diketahui dengan berbagai cara atau metode. Salah satu cara atau metodenya adalah berupa pertanyaan yang mengandung sifat-sifat dalam bentuk skala mulai dari setuju sampai tidak setuju. Tetapi cara ini memiliki kelemahan karena memiliki hasil prediksi yang belum sepenuhnya valid.

Dalam penelitian ini menggunakan model klasifikasi kepribadian berlandaskan dari status media sosial facebook dengan menggunakan Linguistic Inquiry and Word Count (LIWC) dan memakai metode Support Vector Machine. Dengan menggunakan 3 jenis parameter untuk proses seleksi kepribadian antara lain positive emotion, negative emotion dan social word.

Linguistic Inquiry and Word Count (LIWC) adalah program analisis teks psikologi dalam humaniora digital yang dikembangkan oleh James W. Pennebaker, Martha E. Francis pada tahun 1990, LIWC sendiri bertujuan untuk menganalisa psikologi dalam teks atau kalimat digital, Secara tradisional, untuk memahami psikologi dilakukan melalui interpretasi subjektif, analisis kualitatif, dan spekulasi.

Support Vector Machine adalah algoritma untuk meng-klasifikasi data linear dan data non-linear.

Support Vector Machine menerapkan pemetaan non-linear untuk mentransformasi data training awal ke dimensi yang lebih tinggi.

Tujuan dari penelitian tugas akhir ini adalah untuk mengoptimasi nilai tingkat akurasi klasifikasi kepribadian berlandaskan dari status media sosial facebook dengan menerapkan metode Support Vector Machine yang perhitungan

tingkat akurasinya dinyatakan dalam bentuk persentase.

2. METODOLOGI PENELITIAN

2.1. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data. Data yang dimaksud adalah data status dari pengguna sosial media facebook. Pada penelitian ini, data-data status pengguna facebook diambil dengan scraping dengan menggunakan facebook for developer.

2.2. Pengolahan Data

Pada tahap ini, data yang telah dikumpulkan, selanjutnya diolah menggunakan Linguistic Inquiry and Word Count (LIWC), agar didapatkan nilai-nilai untuk mengukur kepribadian pada status pengguna sosial media facebook. data nilai yang didapatkan dari proses menggunakan LIWC selanjutnya diproses menggunakan metode Support Vector Machine untuk menentukan klasifikasinya.

2.3. Perancangan Antarmuka

Pada tahap ini, dilakukan perancangan *user interface* atau antar muka sistem. Antar muka yang akan dirancang menggunakan program matlab.

2.4. Perancangan Arsitektur Sistem

Pada tahap ini, dilakukan perancangan dan penyusunan kerangka berpikir yang menjabarkan bagaimana system bekerja. Rancangan sistem menggambarkan bagaimana system berjalan melalui diagram *use case*.

2.5. Implementasi Sistem

Pada tahap ini, dilakukan pembangunan sistem berdasarkan perancangan yang telah dibuat. Dokumentasi pembangunan sistem akan dilampirkan juga, agar memudahkan dalam penelitian lanjutan di masa yang akan datang.

2.6. Uji Coba Sistem

Pada tahap ini, dilakukan uji coba system terkait klasifikasi kepribadian pengguna facebook menggunakan metode Support Vector Machine..

2.7. Evaluasi

Pada tahap ini, dilakukan evaluasi kinerja system terkait klasifikasi kepribadian pengguna facebook menggunakan metode Support Vector Machine.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Data yang telah dikumpulkan ialah status pengguna facebook dipergunakan dalam proses pengujian sistem.

Tabel 3.1 Dataset status pengguna facebook

Post 1	Post 2	Post 3
(05 November 2011) Call It What You Want~~~~~ Eh Eh, Call It Want You Waaaaannnnntttt ~~~~~	(03 Novemb er 2011) so who has finish the black scholes part???? ??	(01 November 2011) Final Sem seems like a real joy to watch..... Hmhhh m....

Tabel 3.1 contoh dataset status yang diambil pada pengguna facebook.

3.2. Pengolahan Data

Pada tahap dilakukan pemrosesan dataset menggunakan Linguistic Inquiry and Word Count (LIWC) agar mendapatkan nilai-nilai dari kalimat statusnya. Sehingga diperoleh nilai dari proses LIWC data tersebut kemudian dibagi menjadi 2 bagian, 80% untuk data latih dan 20% dipergunakan menjadi data uji.

Tabel 3.2 Pengolahan data status menggunakan LIWC

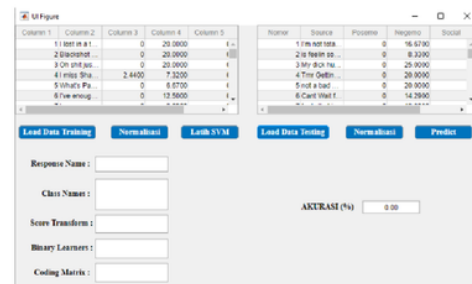
no	Source (A)	posemo	nege mo	social
1	Call It What You Want~~~~~ ~ Eh Eh, Call It Want You Waaaaann nnnnnttt~ ~~~~	0,00	0,00	33,33
2	Women are the joy and pillar of life..!	12,50	0,00	12,50
3	Happy Birthday legend ! ????	33,33	0,00	0,00

Tabel 3.2 di atas menunjukkan nilai dari status pengguna facebook yang telah diproses menggunakan LIWC.

3.3. Perancangan Antarmuka

Sistem dibangun menggunakan aplikasi Matlab. Aplikasi matlab digunakan sebagai program utama sistem, dimana proses pelatihan data latih dan data uji dapat dilakukan menggunakan fitur App Designer.

Berikut merupakan perancangan antar muka aplikasi matlab yang dipakai.



Gambar 3.3. UI Design Matlab

Gambar 3.3 menunjukkan proses pengambilan dataset, Matlab berhasil mengambil data training dan data testing..

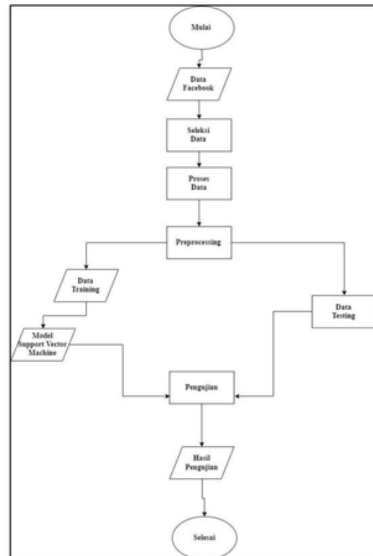
3.4. Perancangan Arsitektur Sistem

Pembuatan arsitektur sistem memuat pola berpikir yang menjabarkan proses sistem menerapkan fungsinya. Sistem yang dibuat untuk penelitian ini dibentuk menggunakan aplikasi matlab.

3.4.1. Flowchart

Flowchart atau diagram alir merupakan diagram yang menggambarkan algoritma, alir kerja, atau proses, yang ditampilkan dalam bentuk simbol-simbol grafis dan panah sebagai penanda urutan. Diagram ini mewakili ilustrasi atau penggambaran penyelesaian masalah. **Flowchart** digunakan untuk menganalisis, mendesain, mendokumentasi, atau memajemen sebuah proses atau program.

Berikut ini **flowchart** yang menjelaskan jalannya proses kerja yang dibuat:



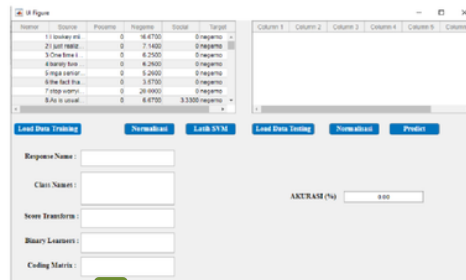
Gambar 3.4 Flowchart Pada Gambar 3.4 menunjukkan **Flowchart**

atau alur proses. Memulai proses, kemudian melakukan pengambilan data dari facebook. Kemudian dilakukan seleksi data yang sesuai dengan criteria remaja usia 20 hingga 30 tahun. Kemudian data hasil seleksi diproses dengan program LIWC untuk mendapatkan nilai dari teks status facebook. Kemudian preprocessing dilakukan normalisasi data, kemudian di bagi menjadi 2 data, 80% untuk data training dan 20% untuk data testing. Kemudian setelah melakukan data training didapatkan model SVM, setelah Model SVM didapatkan dilakukan pengujian. Kemudian dilakukan kroscek hasil pengujian dan keluar nilai hasil akurasi data. Selesai

20

3.5. Proses Pelatihan

Pada tahap ini dilakukan load data training untuk proses pelatihan data. Data latih yang digunakan sebanyak (80%) dari data set.



10

Gambar 3.5 Proses load data training.

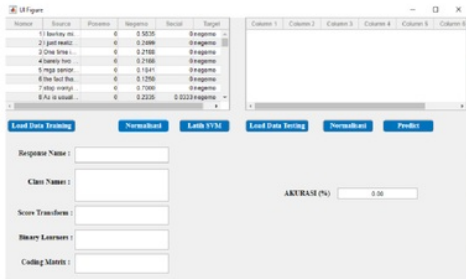
Pada gambar 3.5 menunjukkan proses pengambilan data untuk proses pelatihan.

Setelah data latih ditentukan dilakukan normalisasi data dengan menggunakan metode softmax yang menghasilkan nilai luaran [0.1] dengan rumus :

$$x^j = \frac{x - \min x}{\max x - \min x}$$

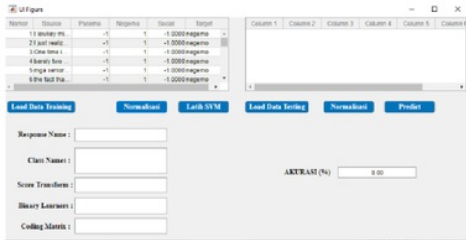
Normalisasi data untuk [-1. 1]

$$x^{II} = 2 \frac{x - \min x}{\max x - \min x}$$



Gambar 3.6 Proses Normalisasi [0.1] data training

Gambar 3.6 menunjukkan proses normalisasi data [0.1] untuk memudahkan atau mempercepat proses pelatihan

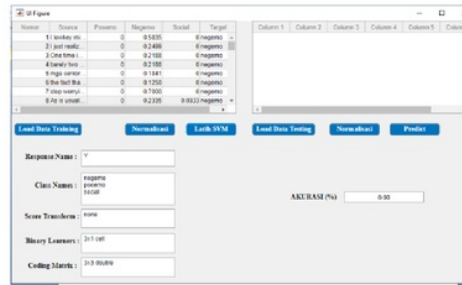


Gambar 3.7 Proses Normalisasi [-1.1] data training

Gambar 3.7 menunjukkan proses normalisasi data [-1.1] untuk memudahkan atau mempercepat proses pelatihan.

Setelah itu dilakukan pelatihan data hingga mendapatkan Model Support Vector Machine yaitu:

1. Response Name : Y.
2. Class Name : negemo (negative emotions), posemo (positive emotions) dan social (social word).
3. Score Transform : none.
4. Binary Learners : 3x1 cell.
5. Coding Matrix : 3x3 double

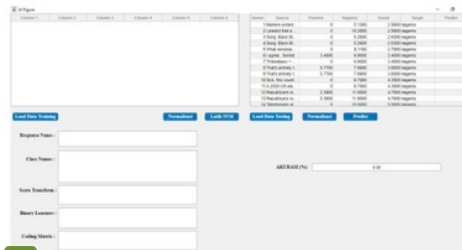


Gambar 3.8 Proses Latih dengan Support Vector Machine (SVM)

Gambar 3.8 menunjukkan proses pelatihan data dengan metode Support Vector Machine.

3.6. Proses Uji

Pada tahap ini dilakukan load data testing untuk proses pengujian data. Data uji yang digunakan sebanyak (20%) dari data set



Gambar 3.9 Proses load data uji

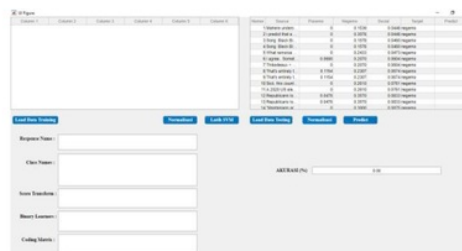
Gambar 3.9 menunjukkan proses pengambilan data untuk data testing atau pengujian data.

Setelah data uji ditentukan dilakukan normalisasi menggunakan metode softmax yang menghasilkan nilai luaran [0-1] dengan rumus :

$$x^I = \frac{x - \min x}{\max x - \min x}$$

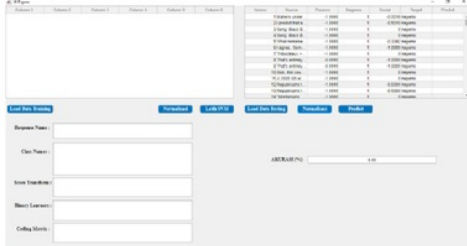
Normalisasi data untuk [-1. 1]

$$x^{II} = 2 \frac{x - \min x}{\max x - \min x}$$



Gambar 3.10 Normalisasi [0.1] data uji

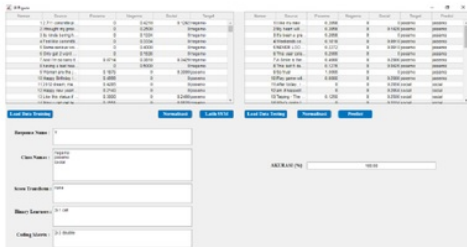
Gambar 3.10 menunjukkan proses normalisasi data [0.1] untuk memudahkan atau mempercepat proses pengujian.



Gambar 3.11 Normalisasi [-1.1] data testing

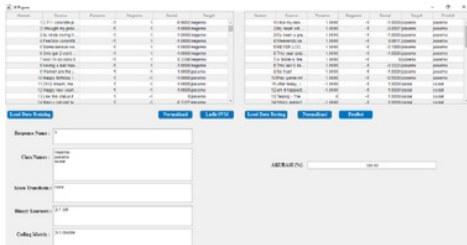
Gambar 3.11 menunjukkan proses normalisasi data [-1.1] untuk memudahkan atau mempercepat proses pengujian

Setelah data diuji berdasarkan hasil dari Model Support Vector Machine yang sudah dibuat hingga didapatkan hasil 100% dan 100% pada pengujian pertama memakai normalisasi [0.1], [-1.1] dan pengujian kedua didapatkan hasil 70.29% dan 99.3% memakai normalisasi [0.1], [-1.1]. Hasil predict bernilai 1 jika benar dan bernilai 0 jika salah



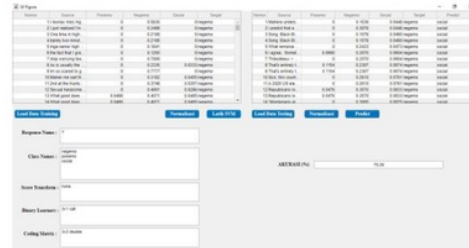
Gambar 3.12. Proses predict data set facebook

Gambar 3.12 menunjukkan proses prediksi data, menggunakan data set dari facebook dan normalisasi [0.1].



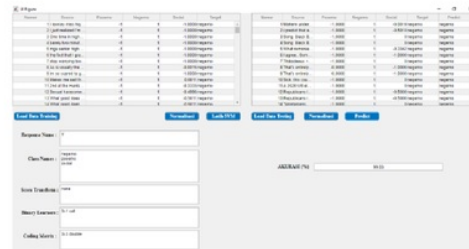
Gambar 3.13 Proses predict data set facebook

Gambar 3.13 menunjukkan proses prediksi data, menggunakan data set dari facebook dan normalisasi [-1.1].



Gambar 3.14 Proses predict data set twitter [0.1]

Gambar 3.14 menunjukkan proses prediksi data, menggunakan data set dari twitter dan normalisasi data [0.1].



Gambar 3.15 Proses predict data set twitter [-1.1]

Gambar 3.15 menunjukkan proses prediksi data, menggunakan data set dari twitter dan normalisasi data [-1.1].

Tabel 3.3 Pengujian Predict pengguna Facebook

No	Target	Predict	Nilai
1	Posemo	Posemo	1
2	Posemo	Posemo	1
3	Posemo	Posemo	1

Tabel 3.4 Pengujian predict pengguna twitter

No	Target	Predict	Nilai
1.	Negemo	Social	0
2.	Negemo	Social	0
3.	Negemo	Social	0

3.7. Evaluasi

Berdasarkan proses pelatihan dan pengujian aplikasi menggunakan UI Design pada matlab, serta metode Support Vector Machine yang diterapkan dapat diambil review serta evaluasi seperti dibawah ini:

1. Hasil pengujian proses pelatihan data latih aplikasi menunjukkan bahwa keluaran aplikasi sesuai

- dengan kebutuhan fungsional yang diharapkan.
2. Hasil pengujian proses testing data uji aplikasi menunjukkan bahwa keluaran aplikasi sesuai dengan kebutuhan fungsional yang diharapkan.
 3. Hasil uji performa aplikasi dan metode SVM di dalamnya menunjukkan bahwa skema normalisasi mempengaruhi hasil prediksi. Proses prediksi yang menggunakan skema normalisasi [0,1] didapatkan nilai prediksi 70.29% menggunakan 2067 dataset dan 100% nilai akurasi pada saat menggunakan 300 dataset. Dan proses prediksi yang menggunakan skema normalisasi [-1.1] didapatkan nilai prediksi 100% ketika menggunakan 300 dataset dan didapatkan 99.3% ketika menggunakan 2067 dataset. Dari hasil tersebut dapat disimpulkan bahwa nilai prediksi yang paling optimal ketika menggunakan skema normalisasi [-1.1].

4. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Dari hasil implementasi dan evaluasi pada penelitian ini, dapat ditarik beberapa poin kesimpulan sebagai berikut:

1. Penelitian ini menghasilkan Sistem aplikasi pada matlab, yang dapat diterapkan untuk mengidentifikasi kepribadian negative dan positif pada pengguna facebook atau social media lainnya.
2. Proses klasifikasi pada penelitian ini, dikembangkan menggunakan metode SVM (Support Vector Machine). Metode ini diterapkan pada proses pelatihan data latih dari 80% data dan 20% data untuk data uji. Pengujian pertama menggunakan 159 data untuk data latih dan 45 data untuk pengujian. Pada pengujian kedua menggunakan 1.653 data untuk data latih dan 414 data untuk pengujian.
3. Melalui proses ujian pertama yang dilakukan memakai 300 status pengguna Facebook, metode SVM

menunjukkan nilai akurasi sebesar 100%. Dan pada pengujian kedua dilakukan dengan 2067 dataset tweets dari pengguna twitter, metode SVM menunjukkan nilai akurasi sebesar 70.29%.

4. Kinerja yang ditunjukkan oleh metode SVM (Support Vector Machine) dapat memenuhi kebutuhan untuk seleksi pemilihan calon pegawai pada dunia industri. Metode ini menunjukkan performa yang baik karena dapat menghasilkan nilai prediksi di atas 70% nilai akurasi. Pengujian dilakukan dua kali menggunakan normalisasi data [0.1] dan [-1.1]. Pada pengujian menggunakan normalisasi [0.1] pertama didapatkan nilai akurasi 100% dan pada pengujian kedua didapatkan nilai akurasi 70.29%. Pada pengujian menggunakan normalisasi [-1.1] pengujian pertama didapatkan nilai akurasi 100% dan pengujian kedua didapati nilai akurasi 99.3%..

4.2. Saran

Melalui penelitian yang sudah dilaksanakan, aplikasi ini pasti banyak memiliki kekurangan, sehingga peneliti memberikan saran bagi pengembangan selanjutnya sebagai berikut:

1. Diperlukan pengembangan UI aplikasi Identifikasi kepribadian agar lebih menarik minat pengguna.
2. Diperlukan pengembangan system agar nilai akurasi tidak menurun walaupun data set yang dipergunakan lebih banyak.
3. Diperlukan pengembangan UI yang lebih menarik, dan pengembangan system agar performansi akurasi tidak menurun sehingga aplikasi identifikasi kepribadian memiliki performa yang lebih baik.

DAFTAR PUSTAKA

- [1]. Kemas Muslim Lhaksana, Fhira Nhita, Duwi Anggraini. 2017. Klasifikasi Kepribadian Berdasarkan Status Facebook Menggunakan Metode Backpropagation. 4(3): 114.
- [2]. Alvaro Ortigosa, dkk. 2013. Predicting User Personality by Mining Social Interactions in Facebook, ed., Journal of Computer and System Sciences: Universidad Autonoma de Madrid, Madrid, Spain: ELSEVIER Hlm 57-71.
- [3]. Salvatore A. Catanese, dkk. 2011. Crawling Facebook for Social Network Analysis Purposes. 15.
- [4]. Golnoosh Farnadi, dkk. 2013. Recognising Personality Traits Using Facebook Status Updates. 6.
- [5]. Yoram Bachrach, Michal Kosinski, Thore Graepel, Pushmeet Kohli, David Stillwell. 2012. Personality and Patterns of Facebook Usage. 4.
- [6]. Dejan Markovikj, Sonja Gievska, Michal Kosinski, David Stillwell. 2013. Mining Facebook Data for Predictive Personality Modeling. 3.
- [7]. Y B N D Artissa, I Asror, S A Faraby. 2019. Personality Classification based on Facebook status text using Multinomial Naive Bayes method. 9.
- [8]. Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. 2015. The development and psychometric properties of VC2015. 18.
- [9]. Nick Guenther, Matthias Schonlau. 2016. Support Vector Machines 16 4: 917-937
Ryan L. Boyd 2017 Psychological Text Analysis in the Digital Humanities.
- [10]. A. Hernandez-Suarez, G. Sanchez-Perez, K. Toscano-Medina, V. Martinez-Hernandez, V. Sanchez, H. Perez-Meana 2018 A Web Scraping Methodology for Bypassing Twitter API Restrictions. 5.
- [11]. Christoph Bergmeir, Rob J. Hyndman, Bonsoo Koo 2018 A note on the validity of cross-validation for evaluating autoregressive time series prediction.
- [12]. Maria Giulia Marini 2018 Language of Emotions, with a Focus on Shame to Be Defeated: The Flower of Emotions and Their Role in a Post-contemporary Society.

KLASIFIKASI KEPRIBADIAN PENGGUNA FACEBOOK MENGUNAKAN SUPPORT VECTOR MACHINE

ORIGINALITY REPORT

% **15**
SIMILARITY INDEX

% **11**
INTERNET SOURCES

% **7**
PUBLICATIONS

% **11**
STUDENT PAPERS

PRIMARY SOURCES

1 Submitted to Universitas Brawijaya % **2**
Student Paper

2 Submitted to Universitas Jember % **1**
Student Paper

3 iopscience.iop.org % **1**
Internet Source

4 www.aclweb.org % **1**
Internet Source

5 Sumarni Adi, Anggit Dwi Hartanto, Ema Utami, Suwanto Raharjo, Irwan Oyong. "The Effect of Data Acquisition Techniques in Profiling Analysis Based on Twitter", 2019 International Conference on Information and Communications Technology (ICOIACT), 2019 % **1**
Publication

6 www.kuntalkumarpal.in % **1**
Internet Source

7 [afit-r.github.io](https://github.com/afit-r)

Internet Source

% 1

8

ikee.lib.auth.gr

Internet Source

% 1

9

dione.lib.unipi.gr

Internet Source

% 1

10

Submitted to Sultan Agung Islamic University

Student Paper

% 1

11

www.cambridge.org

Internet Source

% 1

12

pmb.untag-sby.ac.id

Internet Source

% 1

13

eprints.uny.ac.id

Internet Source

<% 1

14

idec.industri.ft.uns.ac.id

Internet Source

<% 1

15

Submitted to King's College

Student Paper

<% 1

16

pt.scribd.com

Internet Source

<% 1

17

Submitted to President University

Student Paper

<% 1

18

www.mitpressjournals.org

Internet Source

<% 1

19 issuu.com Internet Source <% 1

20 Submitted to Sriwijaya University Student Paper <% 1

21 aeiwell.blogspot.com Internet Source <% 1

22 journal.unusida.ac.id Internet Source <% 1

23 Submitted to Universitas Dian Nuswantoro Student Paper <% 1

24 Ortigosa, Alvaro, Rosa M. Carro, and José Ignacio Quiroga. "Predicting user personality by mining social interactions in Facebook", Journal of Computer and System Sciences, 2013. Publication <% 1

EXCLUDE QUOTES OFF
EXCLUDE BIBLIOGRAPHY OFF

EXCLUDE MATCHES OFF