Jurnal Sistem Komputer dan Informatika (JSON)
Volume 4, Nomor 2, Desember 2022
e-ISSN 2685-998X
DOI 10.30865/json.v4i2.5345

Personality Detection of Twitter Social Media Users using the Support Vector Machine Method

Salsabila Anza Salasa, Warih Maharani*

School of Computing, Informatics, Telkom University, Bandung, Indonesia
Email: [1]anzasalsabila@student.telkomuniversity.ac.id, [2,*]wmaharani@telkomuniversity.ac.id
Correspondent Author Email: wmaharani@telkomuniversity.ac.id
Submitted: **13/12/2022**; Accepted: **27/12/2022**; Published: **31/12/2022**

Abstract– Personality is a person's psychological tendency to carry out certain social behaviors, whether in the form of feelings, thoughts, attitudes and will or actions. Big Five is the most popular and widely used personality model, therefore this proposal uses the Big Five Personality model. In this technological era, humans interact using social media. One of the fastest growing social media is Twitter. Twitter is a social media used by all groups. Every human being has a different personality. Personality detectors are needed for employee recruitment to dig up information about the personality of prospective employees. So personality detection or BigFive Personality can be done through tweets that are shared on Twitter. With this, it is necessary to detect personality using the Support Vector Machine (SVM) method. From the results of the study, it was found that the maximum performance parameter combination in detecting personality on Twitter users was a combination of Linear parameters and C = 10 which obtained an accuracy of 0.979. The data used is the result of crawling on the Twitter site with 146 user usernames and 38853 tweets.

Keywords: Personality, BigFive; Social Media; Twitter; SVM

1. INTRODUCTION

Personality is a combination of behavior and emotions that depend on the environment and biological causes [1]. Personality is a person's psychological tendency to carry out certain social behaviors, whether in the form of feeling, thinking, behaving, and willing or acting. According to M.a.W Bower, personality is the overall behavior of an individual with a certain tendency system that interacts with a series of instructions. Everyone must have a different personality depending on how the individual carries himself. There are several personality models used in predicting personalities, such as the Big Five Personality, MBTI (Myers Briggs Type Indicator) or DISC (Dominance Influence Steadiness Conscientiousness)[2].

Big Five is the most popular and widely used personality model, therefore this proposal uses the Big Five Personality model. BigFive Personality includes 5 main traits namely Extraversion, Agreeableness, Openness, Conscientiousness, and Neuroticism. Extraversion has a firm nature, is easy to get along with the environment, is active, and tends to pursue joy as well as challenges. Agreeableness has a gentle, cooperative, humble, and kind nature, individuals who have this trait also tend to accept and adhere to culture and norms. Openness is sensitive, open-minded, intellectual, and highly imaginative. Conscientiousness tends to be organized, responsible, and meticulous in everything but also careful. Neuroticism has properties that tend to be anxious about something, easily angry, and feel depressed. People who have this trait also easily feel uneasy [3].

Humans interact through social media in the technological era like today. What is meant by social media is online media where users can easily create ideas, express opinions freely, and create content. Activities in using social media as an online communication network which includes text, images, and videos are very helpful in daily activities such as discussions to doing business.

Twitter is a social media that is used by all groups [4]. According to We Are Social January 2020 data, Twitter is the most popular social media network used by all Indonesian people with around 10.65 million users and is also ranked number one among all social media platforms[5]. On Twitter people often express what they think with a single tweet of no more than 140 characters in real time. This Twitter tweet was used as a dataset and then analyzed. Machine learning or machine learning is a method that is widely used in problem analysis such as depression detection and personality detection [6].

Personality detection in this technological era is urgently needed because there is a lot of information about its users, such as companies that use personality detection to detect the personality of their prospective employees. If personality detection is not carried out, the company may have errors in recruiting its human resources [7].

One method that is widely used is the Support Vector Machine (SVM) method. Based on previous research, which was conducted by Shanti ka Valerin Therik and Erwin Budi Setiawan entitled "Detection of the Personality of the Big Five Twitter Users Using the C4.5 Method", where the highest accuracy was obtained at 76.92% obtained from a comparison of Baseline + LIWC + TF data -IDF and implement the SMOTE method [7]. Research conducted by Maulana Amsor using the Support Vector Machine (SVM) method and the Principal Component Analysis (PCA) method obtained an accuracy of 70.9% [8]. Research conducted by Y B N D Artissa, I Astor, and S A Faraby with the title "Personality Classification based on Facebook Status Text Using Multinomial Naive Bayes Method" obtained an accuracy of only 59.9% [9]. Then the research conducted by Bayu Yudha Pratama and Riyanarto Sarno with the title "Personality Classification Based on Twitter Text Using Naive Bayes, KNN, and SVM" obtained an accuracy of 65% [10]. Another study conducted by Tri Ayu Syifa'ur Rohmah entitled

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 4, Nomor 2, Desember 2022
e-ISSN 2685-998X
DOI 10.30865/json.v4i2.5345

"Personality Detection on Twitter Social Media Using the IndoBERT Method" obtained the highest accuracy of 0.46. This study uses 258 Twitter crawling datasets where the test scenario lies in the preprocessing difference dataset [11]. Based on previous research, this study focuses on identifying personal information about people who use social media sites such as Twitter by using a Support Vector Machine (SVM) because it has optimal computing time [12], as well as optimal parser functions and has high data generalization [13]. Lately, the SVM method has been successfully applied to real-world problem situations. In addition, SVM offers a better solution than other conventional methods [14]. SVM has the advantage that SVM can be widened or extended to solve non-linear problems, that is, linear SVM can be converted into non-linear SVM using the kernel [15]. Another advantage of SVM is that SVM has a theoretical basis that can be analyzed clearly and is not a black box. The next generalization strategy is the ability of the current method to classify certain patterns that do not include the data used in the method. The two main causes of generalized error are errors involving the training set and errors caused by VC (Vapnik-Chervokinensis)[14]. This research will use the SVM method using the Count Vectorizer feature. The purpose of this study is to get the maximum performance on the Support Vector Machine (SVM) kernel model approach in detecting personality on Twitter users and to find out the level of accuracy of each kernel approach used by SVM.

2. RESEARCH METODOLOGY

2.1 Architecture System

In this research, the expected result is the performance of the Support Vector Machine (SVM) method using the Count Vectorizer feature to detect personalities on Twitter social media. This study includes five stages for personality detection using tweets which include data collection, data labeling, data pre-processing, data classification, and data evaluation. The flow diagram of the system in this study can be seen in Figure 1 below

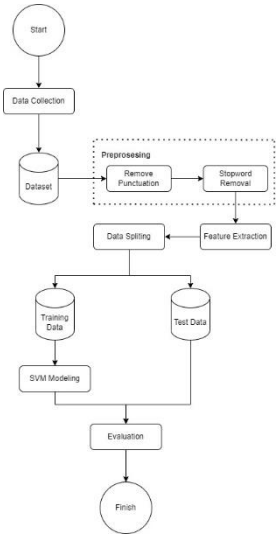

Figure 1. Architecture System

2.2 Data Collection

In this process, the data labeling and data sharing stages are included in BigFive Personality. This stage is carried out by asking respondents to fill out a questionnaire with questions about names, usernames, and personalities that are included in BigFive Personality. 146 names, usernames, and personality data were obtained, the results of the questionnaire are shown in Table 1

Table 1. Data Collection

Name	Username	Personality
Adam	Adamumemo	*Agreebleness*
Citora	Citora	*Conscientiousness*
Clara	Claraang	*Agreebleness*
Danny tandean	Dannytandean	*Agreebleness*
Darna	Darna	*Opennes*
Dzaky	afridzakyy	*Extraversion*
Darana	daranagas	*Neuroticism*
Deria ferlian	deriaferlian	*Agreebleness*
Dimas	dmsim	*Neuroticism*

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 4, Nomor 2, Desember 2022
e-ISSN 2685-998X
DOI 10.30865/json.v4i2.5345

Furthermore, data collection was carried out by crawling data on the Twitter site from each username that had been obtained from 146 questionnaire respondents. The tweet data is an expression in channeling feelings, thoughts, attitudes, and wishes as well as actions that are carried out and contains BigFive Personality which includes 5 main traits, namely Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. 38853 tweets were obtained from the 146 usernames, the results of the crawling data can be seen in Table 2

Table 2. Result Of Crawling Data

Username	Number of Tweet
Adamumemo	*350*
Citora	299
Claraang	289
Dannytandean	250
Darna	197
afridzakyy	178
daranagas	143
deriaferlian	100
dmsim	*78*

2.3 *Pre-processing*

After the data has been successfully crawled, the next step is data preprocessing, data preprocessing to handle unstructured data becomes tidier and cleaner, there are two stages in this process, namely:

a. Remove Punctuation

Remove Punctuation is used to remove punctuation marks, to expedite the SVM training process so that the data is clean, the Remove Punctuation process is shown in table 3

Table 3. Result of remove punctuation

Input	Output
Hari ini aku ketemu mantan :) kenapa dia semakin hari semakin cantik ya??????	Hari ini aku ketemu mantan kenapa dia semakin hari semakin cantik ya

The results of the table above show that the sentences entered still have punctuation marks, and produce output that has had the punctuation removed.

b. Stopword Removal

In this process, removing unnecessary words that are not related to BigFive Personality, and storing important words, the stopwords library used from the literary library, the stopword removal process is shown in table 4

Table 4. Result of stopword removal

Input	Output
“Hari”,”ini”,”aku”,”ketemu”,”mantan”,”kenapa”,”dia”, “semakin”,”hari”,”semakin”,”cantik”,”ya”	“Hari”,“aku”,“ketemu”,“kenapa”,”dia”,“semakin”,“semakin”,”cantik”

The results of the table above show that the sentences entered still contain unnecessary words such as the word "ini" and the word "semakin" that appears twice. Then produce output like the table above.

2.4 *Feature Extraction*

This study uses feature extraction Count Vectorizer where feature extraction is performed by converting text data into vector data. Count Vectorizer is a feature used to count the frequency of words in sentences which are then represented in vector form. The shape of the vector is then saved as a .json file [16]. Which vector represents the count or the number of occurrences of the word? The Count Vectorizer feature extraction stages are shown in the following image example.

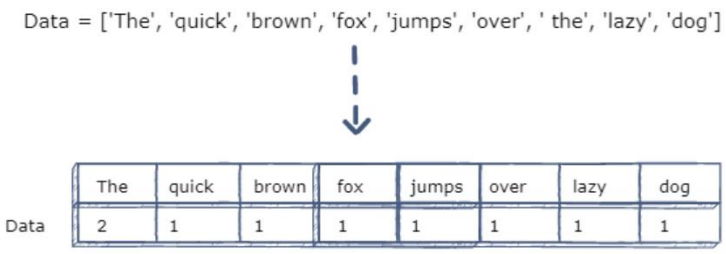

Figure 2. Example of the Count Vectorizer stage

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 4, Nomor 2, Desember 2022
e-ISSN 2685-998X
DOI 10.30865/json.v4i2.5345

In the picture above there is a sentence "The quick brown fox jumps over the lazy dog". The sentence is converted into a list form and then converted into a count form. Then the representation in the example sentence above, namely "The" appears twice, while the other words only appear once.

2.5 *SVM Modelling*

Support Vector Machine (SVM) is a classification method that looks for hyperplanes with the largest margins [17]. Support Vector Machine is similar to ANN which is included in supervised learning, namely finding hyperplanes by separating data sets into two different classes. The hyperplane will find its optimum point when the distance is right in the middle of the class that has been separated [18]. The essence of how the Support Vector Machine (SVM) works is to find the farthest distance from the hyperplane which is divided into two classes. The process of solving the longest distance is repeated several times to find the best hyperplane. Therefore it takes 7 optimizations on SVM to find the maximum distance on the hyperplane with these two classes. There are two forms of optimization intended to find hyperplanes in SVM development, namely Primal Form SVM, and Dual Form SVM[19]. The following is an overview of SVM trying to find the best hyperplane [20].

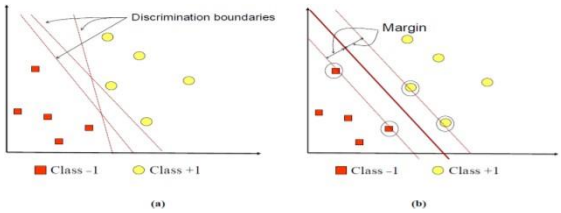

Figure 3. SVM Illustration in finding hyperplanes

Figure 3 SVM Trying to Find a Hyperplane Figure a shows the pattern of two classes, namely +1 and -1. The red color is a pattern belonging to class -1, while the yellow color is a pattern belonging to class +1. The problem with this classification is that SVM tries to find a hyperplane that separates the two classes. Discrimination boundaries are symbols of alternative dividing lines. The best hyperplane can be found by measuring the margins and finding the optimization point. Margin is the distance between the hyperplane and the nearest pattern from each of these two classes. The thickest line in figure b shows the best hyperplane.

At this stage, SVM modeling is carried out. To get the most optimal hyperplane, SVM uses a decision boundary, or is called a decision boundary which will only reveal the specific classifications that exist in the dataset [21]. SVM works based on the Structural Risk Minimization (SRM) principle. The goal is to find a support vector, namely to find a decision rule which is done by selecting a certain subset of the training data (Vapnik, 1988). When applied to classification, the SVM method uses a hyperplane that efficiently extracts two classes from training samples [22]. In the Support Vector Machine method, parameter tuning is used, namely parameters as parameter values or parameter settings to identify and solve problems encountered [23]. One way to solve the SVM problem is to use the following formula:

$$L(\vec{w}, b, \alpha) = \frac{1}{2}\|\vec{w}\|^2 - \sum_{i=1}^{l} \alpha_i (y_i((\vec{x}_i.\vec{w} + b) - 1)) \quad (1)$$

α_i is a Lagrange Multiplier which has a zero or positive value ($\alpha_i \geq 0$). The optimal value or the best value from Equation above can be solved by minimizing the interaction of l with $\vec{w}$, and b, and maximizing the interaction of l with α_i. By understanding the principle that the ideal gradient is equal to zero, the above equation can be modified to a minimum, similar to the equation below.
Maximum:

$$\sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} \alpha_i \alpha_j y_i\, y_j \vec{x}_i.\vec{x}_j \quad (2)$$

Subject to :

$$\alpha_i \geq 0\ (i = 1,2,\dots,l) \sum_{i=1}^{l} \alpha_i y_i = 0 \quad (3)$$

From the equations above we get the result ai which is positive. Data with a positive value is said to be correlated with ai, that is the support vector value [14].

3. RESULT AND DISCUSSION

3.1 Performance Evaluation

In this study using an approach with C parameters and Kernel parameters. For example, parameter C uses the values 01, 1, and 10 and then proceeds with using Kernel parameters, namely Linear, Poly, RBF, and Sigmoid parameters. The kernel is a parameter that can be used to classify non-linear data [24].

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 4, Nomor 2, Desember 2022
e-ISSN 2685-998X
DOI 10.30865/json.v4i2.5345

As a result, SVM separates the data set into two different classes, then from these two classes the optimal point will be determined and the farthest distance from the optimal point or hyperplane will be determined to predict. Then the highest accuracy value is chosen between these optimal points as the final prediction result.

3.1.1 Parameter C = 0.1

In parameter C = 0.1, an accuracy of 0.93 is obtained. Based on the results of the personality detectors produced, which are included in the BigFive Personality, as much as 20.73% are included in Agreeableness, 16.74% are included in Conscientiousness, 22.55% are included in Extraversion, 15.19% are included in Neuroticism, and 18.15% are included in Openness. The evaluation matrix for using the parameter C = 0.1 is shown in Figure 4.

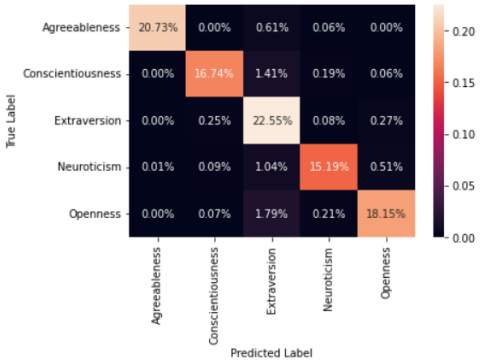

Figure 4. The evaluation matrix for the use of parameter C = 0.1

3.1.2 Parameter C = 1

In parameter C = 1, an accuracy of 0.97 is obtained. Based on the results of the personality detectors produced, which are included in the BigFive Personality, 21.11% are included in Agreeableness, 17.76% are included in Conscientiousness, 22.92% are included in Extraversion, 16.30% are included in Neuroticism, and 19.21% are included in Openness. The evaluation matrix for using the parameter C = 1 is shown in Figure 5.

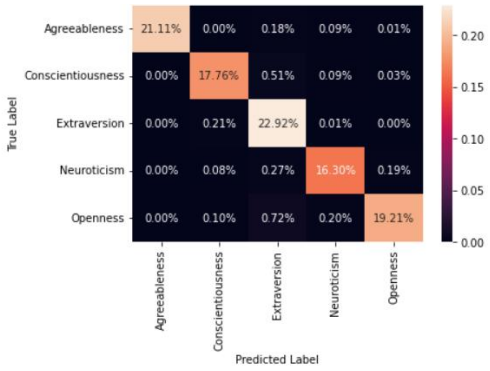

Figure 5. The evaluation matrix for the use parameter C = 1

3.1.3 Parameter C = 10

In parameter C = 10, an accuracy of 0.97 is obtained. Based on the results of the resulting personality detector, which is included in BigFive Personality, 21.12% is included in Agreeableness, 17.76% is included in Conscientiousness, 22.80% is included in Extraversion, 16.36% is included in Neuroticism, and 19.50% included in Opennes. The evaluation matrix for using the parameter C = 10 is shown in Figure 6.

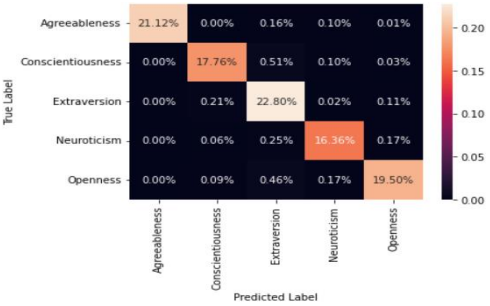

Figure 6. The evaluation matrix for the use parameter C = 10

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 4, Nomor 2, Desember 2022
e-ISSN 2685-998X
DOI 10.30865/json.v4i2.5345

After experimenting with the C parameter approach, the accuracy of each parameter has a different value, the accuracy for parameters C = 1 and C = 10 is almost the same, only parameter C = 10 gets the highest accuracy, the difference is in the prediction of Neuroticism and Openness personality, in parameter C = 1 obtained Neuroticism of 16.30% and Openness of 19.21% while in parameter C = 10 obtained Neuroticism of 16.36% and Openness of 19.50% The results of the accuracy of each parameter are shown in the following table.

Table 5. The results of the accuracy of each parameter

Parameter	*Agreebleness*	*Conscientiousness*	*Extraversion*	*Neuroticism*	*Opennes*	Akurasi
C = 01	20.73%	16.74%	22.55%	15.19%	18.15%	0.93
C = 1	21.11%	17.76%	22.92%	16.30%	19.21%	0.972
C = 10	21.12%	17.76%	22.80%	16.36%	19.50%	0.975

From the results of the C parameter experiment, there was no significant difference, almost all experiments obtained the same accuracy value, but in this study, only 1 parameter was used, and the highest accuracy was obtained, namely C = 10, so this parameter will be used in the next process. and combined with kernel parameters. Next is the Kernel parameter which consists of Linear, Poly, RBF, and Sigmoid parameters.

3.1.4 Parameter Kernel Linier

In the linear parameter combined with the parameter C = 10, an even higher accuracy is obtained, namely 0.979. Based on the resulting personality detector results, which are included in the BigFive Personality, as much as 21.35% are included in Agreeableness, 18.21% are included in Conscientiousness, 22.43% are included in Extraversion, 16.21% belong to Neuroticism, and 19.72% belongs to Openness. Evaluation matrix on the use of combined linear kernel parameters C = 10 is shown in Figure 7.

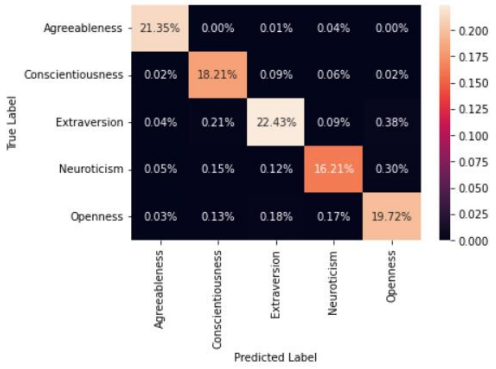

Figure 7. The evaluation matrix uses the combined linear kernel parameters C = 10

3.1.5 Parameter Kernel Poly

In the poly parameter combined with the parameter C = 10, a lower accuracy is obtained, namely 0.976. Based on the resulting personality detector results, which are included in the BigFive Personality, as much as 21.40% are included in Agreeableness, 18.38% are included in Conscientiousness, 22.32% are included in Extraversion, 15.64% are included in Neuroticism, 19.95% are included in Openness. The evaluation matrix for the use of the combined poly kernel parameter C = 10 is shown in Figure 8.

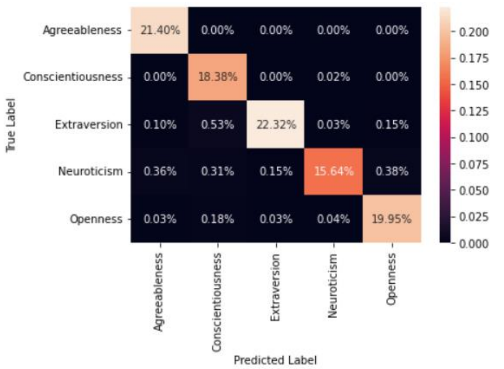

Figure 8. The evaluation matrix uses the combined poly kernel parameters C = 10

3.1.6 Parameter Kernel RBF

In the RBF parameter combined with the parameter C = 10, a slightly lower accuracy is obtained, namely 0.975. Based on the resulting personality detector results, which are included in the BigFive Personality, as much as

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 4, Nomor 2, Desember 2022
e-ISSN 2685-998X
DOI 10.30865/json.v4i2.5345

21.12% are included in Agreeableness, 17.76% are included in Conscientiousness, 22.80% are included in Extraversion, 16.36% belong to Neuroticism, and 19.50% belongs to Openness. The evaluation matrix for using the RBF kernel parameters combined with C = 10 is shown in Figure 9.

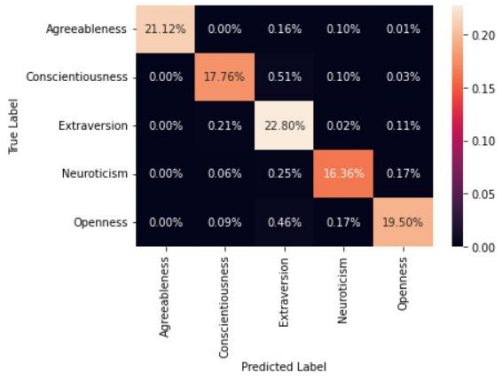

Figure 9. The evaluation matrix uses the combined RBF kernel parameters C = 10

3.1.7 Parameter Kernel Sigmoid

In the Sigmoid parameter combined with the parameter C = 10, a lower accuracy is obtained, which is 0.970. Based on the results of the resulting personality detector, which is included in the BigFive Personality, as much as 21.24% is included in Agreeableness, 17.98% is included in Conscientiousness, 22.26% is included in Extraversion, 16.13% are included in Neuroticism, and 19.44% are included in Openness. The evaluation matrix for using the RBF kernel parameters combined with C = 10 is shown in Figure 10.

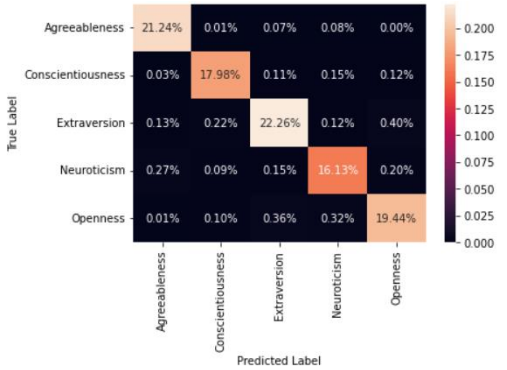

Figure 10. The evaluation matrix uses the combined sigmoid kernel parameters C = 10

After conducting experiments with a combination of C and Kernel parameters, the accuracy of each combination of parameters gets almost the same accuracy, but there are still differences. The results of the parameter combination experiments are shown in the following table.

Table 6. The results of the parameter combination experiments

Parameter	*Agreebleness*	*Conscientiousness*	*Extraversion*	*Neuroticism*	*Opennes*	Akurasi
Linear + C = 10	21.35%	18.21%	22.43%	16.21%	19.72%	0.979
Poly + C = 10	21.40%	18.38%	22.32%	15.64%	19.95%	0.976
RBF + C = 10	21.12%	17.76%	22.80%	16.36%	19.50%	0.975
Sigmoid + C= 10	21.24%	17.98%	22.26%	16.13%	19.44%	0.970

From the results of the parameter combination experiments, it was found that the parameter combination obtained the highest accuracy, namely the combination of Linear parameters and C = 10 with an accuracy value of 0.979. The personality prediction results from 146 usernames and 38853 tweet data using the Support Vector Machine (SVM) method are shown in the following table.

Table 7. The personality prediction results from 146 usernames and 38853 tweet data

Username	**Personality**
Adamumemo	*Agreebleness*
Citora	*Conscientiousness*
Claraang	*Agreebleness*
Dannytandean	*Agreebleness*

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 4, Nomor 2, Desember 2022
e-ISSN 2685-998X
DOI 10.30865/json.v4i2.5345

Username	Personality
Darna	*Opennes*
afridzakyy	*Extraversion*
daranagas	*Neuroticism*
deriaferlian	*Agreebleness*
dmsim	*Neuroticism*

4. CONCLUSION

From the results of the research above, it was found that the combination of parameters for maximum performance in personality detection on Twitter users used 146 user usernames and 38853 tweets using the Support Vector Machine (SVM) method, namely a combination of Linear and C=10 parameters which obtained an accuracy of 0.979. There is no significant difference in each experiment, these results indicate that the Support Vector Machine (SVM) is very suitable for personality detection. In this study the authors propose the Support Vector Machine (SVM) method for use in personality detection, using a combination of Linear parameters and C = 10 using more datasets, and more than 146 usernames.

REFERENCES

[1] G. A. Buntoro, "Analisis Sentimen Hatespeech Pada Twitter Dengan Metode Naïve Bayes Classifier Dan Support Vector Machine," *J. Din. Inform.*, vol. 5, no. 2, pp. 1–13, 2016.
[2] T. Tandera, Hendro, D. Suhartono, R. Wongso, and Y. L. Prasetio, "Personality Prediction System from Facebook Users," *Procedia Comput. Sci.*, vol. 116, pp. 604–611, 2017, doi: 10.1016/j.procs.2017.10.016.
[3] S. Roccas, L. Sagiv, S. H. Schwartz, and A. Knafo, "The Big Five personality factors and personal values," *Personal. Soc. Psychol. Bull.*, vol. 28, no. 6, pp. 789–801, 2002, doi: 10.1177/0146167202289008.
[4] D. Alita, Y. Fernando, and H. Sulistiani, "Implementasi Algoritma Multiclass Svm Pada Opini Publik Berbahasa Indonesia Di Twitter," *J. Tekno Kompak*, vol. 14, no. 2, p. 86, 2020, doi: 10.33365/jtk.v14i2.792.
[5] Alkatiri Bin Muhammad Awad, Nadiah Zhafira, and S. Nasution Nada Adinda, "Opini Publik Terhadap Penerapan New Normal Di Media Sosial Twitter," *Cover. J. Strateg. Commun.*, vol. 11, no. 1, 2020, [Online]. Available: www.covid19.go.id.
[6] D. Markovikj, S. Gievska, M. Kosinski, and D. Stillwell, "Mining facebook data for predictive personality modeling," *AAAI Work. - Tech. Rep.*, vol. WS-13-01, pp. 23–26, 2013, doi: 10.1609/icwsm.v7i2.14466.
[7] S. V. Therik, "Deteksi Kepribadian Big Five Pengguna Twitter Dengan Metode C4.5," vol. 8, no. 5, pp. 10277–10287, 2021.
[8] J. Ilmiah, I. Komputa, B. P. Utama, K. K. Purnamasari, and J. D. Bandung, "Support Vector Machine Dalam Sistem Pendeteksi Kepribadian Berdasarkan Pola Tanda Tangan," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, 2018.
[9] Y. B. N. D. Artissa, I. Asror, and S. A. Faraby, "Personality Classification based on Facebook status text using Multinomial Naïve Bayes method," *J. Phys. Conf. Ser.*, vol. 1192, no. 1, 2019, doi: 10.1088/1742-6596/1192/1/012003.
[10] B. Y. Pratama and R. Sarno, "Personality classification based on Twitter text using Naive Bayes, KNN and SVM," *Proc. 2015 Int. Conf. Data Softw. Eng. ICODSE 2015*, pp. 170–174, 2016, doi: 10.1109/ICODSE.2015.7436992.
[11] G. Informatika, S. Bits, S. H. Issn, T. Ayu, and W. Maharani, "Deteksi Kepribadian di Media Sosial Twitter Menggunakan Metode IndoBERT," vol. 4, no. 2, pp. 448–453, 2022, doi: 10.47065/bits.v4i2.1895.
[12] E. Ronando, M. I. Irawan, and E. Apriliani, "A Hybrid Approach Support Vector Machine (SVM) - Neuro Fuzzy For Fast Data Classification," *IPTEK J. Proceeding Ser.*, vol. 1, pp. 105–106, 2015.
[13] F. Damayanti, A. Z. Arifin, and R. Soelaiman, "Pengenalan Citra Wajah Menggunakan Metode Two-Dimensional Linear Discriminant," vol. 5, no. 3, pp. 147–156, 2010.
[14] D. A. Pisner and D. M. Schnyer, "Support Vector Machine - Teori dan Aplikasinya dalam Bioinformatika," *Mach. Learn. Methods Appl. to Brain Disord.*, pp. 101–121, 2019, doi: 10.1016/B978-0-12-815739-8.00006-7.
[15] "Penerapan Metode Support Vector Machine (SVM) untuk Mendeteksi Penyalahgunaan Narkoba.pdf." .
[16] M. Priandi and Painem, "Analisis Sentimen Masyarakat Terhadap Pembelajaran Daring di Era Pandemi Covid-19 pada Media Sosial Twitter Menggunakan Ekstraksi Fitur Countvectorizer dan Algoritma K-Nearest Neighbor," *Semin. Nas. Mhs. Ilmu Komput. dan Apl. Jakarta-Indonesia*, no. September, pp. 311–319, 2021.
[17] H. N. Irmanda and Ria Astriratma, "Klasifikasi Jenis Pantun Dengan Metode Support Vector Machines (SVM)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 5, pp. 915–922, 2020, doi: 10.29207/resti.v4i5.2313.
[18] R. Between, S. J. Impact, T. Related, E. Indexes, and Z. Lian-feng, "Perbandingan Metode Decision Tree dan Support Vector Machine untuk Analisis Sentimen pada Instagram Mengenai Kinerja PSSI," vol. 7, no. 3, pp. 9936–9948, 2020.
[19] F. S. Jumeilah, "Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 1, pp. 19–25, 2017, doi: 10.29207/resti.v1i1.11.
[20] I. C. R. Drajana, "Metode Support Vector Machine Dan Forward Selection Prediksi Pembayaran Pembelian Bahan Baku Kopra," *Ilk. J. Ilm.*, vol. 9, no. 2, pp. 116–123, 2017, doi: 10.33096/ilkom.v9i2.134.116-123.
[21] L. Mutawalli, M. T. A. Zaen, and W. Bagye, "KLASIFIKASI TEKS SOSIAL MEDIA TWITTER MENGGUNAKAN SUPPORT VECTOR MACHINE (Studi Kasus Penusukan Wiranto)," *J. Inform. dan Rekayasa Elektron.*, vol. 2, no. 2, p. 43, 2019, doi: 10.36595/jire.v2i2.117.
[22] D. S. Ramdan, "Komparasi Antara Algoritma Support Vector Machine Dan Artificial Neural Network Untuk Pemodelan Data Time Series," *Tedc*, vol. 8, no. 2, pp. 154–159, 2014.
[23] C. Huang, Y. Li, and X. Yao, "A Survey of Automatic Parameter Tuning Methods for Metaheuristics," *IEEE Trans. Evol. Comput.*, vol. 24, no. 2, pp. 201–216, 2020, doi: 10.1109/TEVC.2019.2921598.

Jurnal Sistem Komputer dan Informatika (JSON)
Volume 4, Nomor 2, Desember 2022
e-ISSN 2685-998X
DOI 10.30865/json.v4i2.5345

[24] M. . Imelda A.Muis & Muhammad Affandes, "Penerapan Metode Support Vector Machine (SVM) Menggunakan Kernel Radial Basis Function (RBF) Pada Klasifikasi Tweet," *Sains, Teknol. dan Ind. Sultan Syarif Kasim Riau*, vol. 12, no. 2, pp. 189–197, 2015.