

**DETECTION OF TEXTUAL CYBERBULLYING BY
USING SUPPORT VECTOR MACHINE (SVM)**

SKRIPSI



Oleh :
ROCHMAD ADHIM
NIM. 15650107

**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2019**

**DETECTION OF TEXTUAL CYBERBULLYING BY
USING SUPPORT VECTOR MACHINE (SVM)**

SKRIPSI

**Diajukan kepada:
Fakultas Sains dan Teknologi
Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)**

**Oleh:
ROCHMAD ADHIM
NIM. 15650107**

**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2019**

LEMBAR PERSETUJUAN
DETECTION OF TEXTUAL CYBERBULLYING BY
USING SUPPORT VECTOR MACHINE (SVM)


SKRIPSI


Oleh :
ROCHMAD ADHIM
NIM. 15650107

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal : 21 Juni 2019

Dosen Pembimbing I

Dosen Pembimbing II


Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008


Roro Inda Melani, M.T., M.Sc
NIP. 19780925 200501 2 008

Mengetahui,
Ketua Jurusan Teknik Informatika
Jurusan Teknik Informatika Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang




Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

LEMBAR PENGESAHAN
DETECTION OF TEXTUAL CYBERBULLYING BY
USING SUPPORT VECTOR MACHINE (SVM)

SKRIPSI





Oleh :
ROCHMAD ADHIM
NIM. 15650107

Telah Dipertahankan di Depan Dewan Penguji
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Pada Tanggal 21 Juni 2019

Susunan Dewan Penguji

- | | | |
|-----------------------|---|---|
| 1. Penguji Utama | : | <u>Irwan Budi Santoso, M.Kom</u>
NIP. 19770103 201101 1 004 |
| 2. Ketua Penguji | : | <u>Dr. M. Amin Hariyadi, M.T</u>
NIP. 19670118 20050101 1 001 |
| 3. Sekretaris Penguji | : | <u>Dr. Cahyo Crysdian</u>
NIP. 19740424 200901 1 008 |
| 4. Anggota Penguji | : | <u>Roro Inda Melani, M.T., M.Sc</u>
NIP. 19780925 200501 2 008 |

Tanda tangan

()
()
()
()

Mengetahui,
Ketua Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : Rochmad Adhim
NIM : 15650107
Jurusan : Teknik Informatika
Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-banar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan Skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 21 Juni 2019

Yang membuat pernyataan,



MOTTO

“

DOA

ADALAH SATU-SATUNYA
KECURANGAN YANG DIANJURKAN

”



HALAMAN PERSEMBAHAN

الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ

“ Skripsi ini saya persembahkan untuk semua yang saya kenal, baik yang selalu mendukung saya maupun yang selalu memberikan saya pelajaran dan pemahaman. ”



KATA PENGANTAR

Assalamualaikum Wr. Wb.

Puji syukur kehadirat Allah SWT atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan studi dengan tepat waktu sekaligus menyelesaikan skripsi dengan baik dan lancar. Tujuan dari penyusunan skripsi ini guna memenuhi salah satu syarat untuk bisa menempuh ujian sarjana komputer pada Fakultas Sains dan Teknologi (FSAINTEK) Program Studi Teknik Informatika di Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang. Didalam pengerjaan skripsi ini telah melibatkan banyak pihak yang sangat membantu dalam banyak hal. Oleh sebab itu, disini penulis sampaikan rasa terima kasih sedalam-dalamnya kepada:

1. Prof. Dr. Abdul Haris, M.Ag, selaku Rektor Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
2. Dr. Sri Harini, M.Si, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
3. Dr. Cahyo Crysdiyan, selaku Ketua Jurusan Teknik Informatika sekaligus Dosen Pembimbing I yang telah membimbing dalam penyusunan skripsi ini hingga selesai.
4. Roro Inda Melani, M.T., M.Sc, selaku Dosen Pembimbing II yang telah membimbing dalam penyusunan skripsi ini hingga selesai.
5. Irwan Budi Santoso, M.Kom dan Dr, Amin Hariyadi, M.T, Selaku Dosen penguji yang telah memberikan banyak saran untuk kebaikan penulis.

6. Ayah, Ibu, dan Kakak tercinta yang telah banyak memberikan doa dan dukungan kepada penulis secara moril maupun materil hingga skripsi ini dapat terselesaikan.
7. Teman-teman Teknik Informatika Interface 2015 yang senantiasa memberi motivasi dan berjuang bersama selama menjadi mahasiswa.
8. Semua pihak yang telah banyak membantu dalam penyusunan skripsi ini yang tidak bisa penulis sebutkan semuanya.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat kekurangan dan penulis berharap semoga skripsi ini bisa memberikan manfaat kepada para pembaca khususnya bagi penulis secara pribadi.

Malang, 21 Juni 2019

Penulis

LIST OF CONTENT

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
PERNYATAAN KEASLIAN TULISAN	iii
MOTTO	v
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR.....	vii
LIST OF CONTENT	ix
LIST OF FIGURE	ix
LIST OF TABLE	ixi
ABSTRAK	xii
ABSTRACT	xiii
ملخص.....	xiv
CHAPTER I INTRODUCTION.....	1
1.1 Research Background	1
1.2 Research Question	4
1.3 Research Objectives	4
1.4 Research Scope.....	5
1.5 Research Benefit.....	5
CHAPTER II LITERATUR REVIEW.....	6
2.1 Detection of Cyberbully	6
2.2 Support Vector Machine	11
CHAPTER III SYSTEM DESIGN AND IMPLEMENTATION.....	17
3.1 Data Collection and Analysis	17
3.2 System Design and Implementation	18
3.2.1 Input	19
3.2.2 Preprocessing	19
3.2.3 Feature Extraction.....	21
3.2.4 Training.....	23
3.2.5 Testing	26
3.2.6 Data Crawl	27
3.2.7 Database Design	28
3.2.8 Input Data Implementation	29
3.2.9 Generating Dictionary.....	31
3.2.10 Preprocessing Implementation.....	32
3.2.11 Get Feature of Text (Exctraction of Feature).....	37
3.2.12 Training Implementation	39
3.2.13 Testing Implementation	52
3.2.14 Design Interface Impelementation.....	55
CHAPTER IV RESULT AND ANALYSIS.....	61
4.1 Experimental Setup	61
4.2 Experimental Result	63
4.3 Analysis and Integration.....	79
CHAPTER V CONCLUSIONS AND SUGGESTIONS.....	83
REFERENCE	84

LIST OF FIGURE

Figure 3.1 Data Collection Flow	18
Figure 3.2 System Design	19
Figure 3.3 Text Cleaning Design	20
Figure 3.4 Preprocessing Design.....	21
Figure 3.5 Feature Extraction Design	22
Figure 3.6 Training Process Flowchart	24
Figure 3.7 Testing Process Flowchart	27
Figure 3.9 Communication Data	30
Figure 3.10 Generating Dictionary Script	32
Figure 3.11 Data Before Preprocessing.....	33
Figure 3.12 Data After Preprocessing	33
Figure 3.13 Preprocessing Script	34
Figure 3.14 Example Cleaning in Text	34
Figure 3.15 Cleaning Script Code	35
Figure 3.16 Example Stopword in Process	35
Figure 3.17 Stopword Script Code.....	36
Figure 3.18 Example Second Cleaning in Text	36
Figure 3.19 Second Cleaning Script Code	36
Figure 3.20 Example Get Feature of Text.....	37
Figure 3.21 Get Feature Script Code.....	38
Figure 3.22 Initiate First Weight Script Code.....	40
Figure 3.23 Updating Weight Script Code.....	40
Figure 3.24 Training Script Code Part.1	51
Figure 3.25 Training Script Code Part.2	52
Figure 3.26 Testing Script Code Part.2	55
Figure 3.27 Display of Home Page	56
Figure 3.28 Display of Home Page with Result.....	56
Figure 3.29 Display of Login Page	57
Figure 3.30 Display of Register Page.....	57
Figure 3.31 Display of Training Page	58
Figure 3.32 Display of Testing Page	59
Figure 3.33 Display of Docs Page.....	59
Figure 3.34 Display of Dictionary Page.....	60
Figure 4.1 Testing Data Chart	70
Figure 4.2 Testing Result Chart	74

LIST OF TABLE

Table 3.1 List Feature.....	22
Table 3.2 Database data.....	29
Table 3.3 Dictionary Table.....	31
Table 3.4 Training Data.....	41
Table 3.5 Example Feature Extraction	49
Table 4.1 Confussion Matrix	61
Table 4.2 Data Testing	63
Table 4.3 Testing Result.....	63
Table 4.4 Result Testing Comparison	70
Table 4.5 Accuration Result.....	74
Table 4.6 Speed Testing With 40 Text.....	76
Table 4.7 Testing Speed With 72 Text.....	76
Table 4.8 Testing Speed With 151 Text.....	77
Table 4.9 Data for Measurement Reliability	77
Table 4.10 False Detection Data Sample	779

ABSTRAK

Adhim, Rochmad. 2019. **Detection Of Textual Cyberbullying By Using Support Vector Machine (SVM)**. Skripsi. Jurusan Teknik Informatika. Fakultas Sains dan Teknologi. Universitas Islam Negeri Maulana Malik Ibrahim Malang. .

Pembimbing: (i) Dr. Cahyo Crys dian

(ii) Roro In da Melani, MT., M.Sc

Kata Kunci : Cyberbullying, Detection, SVM

Cyberbullying adalah masalah yang kerap terjadi dalam interaksi di social media. Salah satu yang menjadi tantangan setiap developer social media adalah mengatasi tindakan cyberbullying. Salah satu cara yang bisa dilakukan adalah dengan melakukan deteksi dini terhadap text yang disubmit oleh user dalam percakapan di social media. Mengukur akurasi text cyberbullying menggunakan Support Vector Machine merupakan tujuan dari penelitian ini. Penelitian ini juga menyumbangkan beberapa parameter yang dapat digunakan dalam mendeteksi text cyberbullying. Proses training dan testing menggunakan data yang telah dinilai oleh beberapa mahasiswa. Hasil penelitian menunjukkan nilai akurasi sekitar 70%, nilai presisi sebesar 25%, dan nilai recall sebesar 21%. serta menunjukkan kecepatan yang tidak terlampaui jauh saat menggunakan jumlah data yang berbeda-beda untuk dilakukan proses deteksi. Dan nilai reliability sistem sebesar 60%.

ABSTRACT

Adhim, Rochmad. 2019. **Detection Of Textual Cyberbullying By Using Support Vector Machine (SVM)**. Undergraduate Theses. Informatics Engineering Department. Faculty of Science and Technology. State Islamic University of Maulana Malik Ibrahim Malang.

Advisors: (i) Dr. Cahyo Crysdiان
(ii) Roro Inda Melani, MT., M.Sc

Keywords: Cyberbullying, Detection, SVM

Cyberbullying is a problem that often occurs in interactions on social media. One of the challenges for every developer social media is decrease cyberbullying crime. One way that can be done is by conducting early detection of text submitted by users in conversations on social media. Measuring the accuracy of cyberbullying text using Support Vector Machine is the goal of this study. This research also contributes several parameters that can be used to detect cyberbullying text. The training and testing process uses data that has been assessed by several students. The results showed an accuracy of around 70%, a precision is 25%, and a recall value is 21%. And shows the speed that is not too far away when using different amounts of data to do the detection process. And the value of system reliability is 60% error.

ملخص البحث

عظيم، رحمة ٢٠١٩. الكشف عن التسلط عبر الإنترنت نصية باستخدام آلة الدعم الموجه (SVM). البحث الجامعي. شعبة المعلوماتية. كلية العلوم والتكنولوجيا. الجامعة الإسلامية الحكومية مولانا مالك إبراهيم مالانج.

المشرف: (١) الدكتور جهيو كريسديان
(٢) رورو اند ميلاني، الماجستير

الكلمات الرئيسية: التسلط، الكشف، آلة الدعم الموجه (SVM)

البلطجة الإلكترونية هي مشكلة تحدث غالبًا في التفاعلات على وسائل التواصل الاجتماعي. أحد التحديات التي يواجهها كل مطور وسائط اجتماعية هو التغلب على التسلط عبر الإنترنت. طريقة واحدة للقيام بذلك هي مع الكشف المبكر عن النص المقدم من قبل المستخدم في المحادثة على وسائل الاعلام الاجتماعية. الهدف من هذه الدراسة هو قياس دقة البلطجة الإلكترونية للنصوص باستخدام آلة الدعم الموجه (SVM). هذه الدراسة أيضا التبرع بعض المعلومات التي يمكن استخدامها في الكشف عن التسلط النص. تستخدم عملية التدريب والاختبار البيانات التي تم تقييمها من قبل العديد من الطلاب. أظهرت النتائج دقة حوالي ٧٠ ٪ ، وقيمة دقة ٢٥ ٪ ، وقيمة سحب ٢١ ٪. ويبين السرعة غير البعيدة عند استخدام كميات مختلفة من البيانات للقيام بعملية الكشف. وقيمة موثوقية النظام هي ٦٠ ٪.

CHAPTER I

INTRODUCTION

1.1 Research Background

Cyberbully is the representation of bully in online social media. Patchin & Hinduja (2012) defines cyberbully as an intentional activity carried out repeatedly which poses a danger that is carried out through the use of computers, cellphones, and other electronic devices. Mostly, cyberbully is carried out with the aim of threatening, humiliating other people, generally occurs in various media in cyberspace. Cyberbully is the most frequent happen in account that is categorized as an controversial person, but is possible if cyberbully happening to normal people. The number of victims who experienced cyberbully mentioned by Patchin & Hinduja (2012) began around 10-40% more. Many forms of cyberbullying, including harassment, threats, exclusion, etc. One form of cyberbully action is to express openly dislike to someone about someone's potential or actions. Cyberbully can attack anyone, as evidenced by around 11 percent of respondents in Patchin's study experiencing cyberbully. Generally cyberbully victims around the age of adolescents to adulthood.

Cyberbully is considered more cruel and dangerous than bullying that occurs in the real life. The reason is first, the identity of cyberbully perpetrators who can use anonymously makes it difficult to find the culprit. second, cyberbully coverage is broader than bullying in the real life, meaning that if bullying in real life is only done by someone who is known, unlike cyberbully, which can be done by everyone who knows or not, both there is a connection and is not related at all. Third,

cyberbully acts like virus, when a number of people simultaneously attack, some groups join in the attack too, without knowing the truth behind it. Fourth, the absence of pity will arise on the target because the perpetrators cannot see the expressions of the victims directly, and consider their actions to be right. Fifth, cyberbully is not known to their closest people of them, such as parents, because teenagers tend to cover their problem, as Juvonen (2008) explained that teenagers are reluctant to tell their parents about the online incidents that happened to them because they did not want parents to limit their online activities.

Cyberbully in terms of law is an act that violates. Both in Indonesia and in the world have banned cyberbully actions. Forty-three states in the US currently have intimidation prevention laws as of June 2010 (Bullypolice.org, 2010). At least 21 states have laws against cyberbully (Brookover, 2008; Bullypolice.org, 2010). These states are Arkansas, California, Delaware, Florida, Georgia, Idaho, Iowa, Kansas, Maryland, Minnesota, Missouri, Nebraska, New Hampshire, New Jersey, North Carolina, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, and Washington. (Brookover, 2008; Bullypolice.org, 2010). Like the other Country, Indonesia has a law as a form of resistance to cyberbully's crimes. Indonesia decreases the level of cyberbully by making an ITE law that requires punishment for anyone who does cyberbully in any form. There is a sound from the law which is Article 27 paragraph 1,3,4 which reads:

“Any Person who knowingly and without authority distributes and/or transmits and/or causes to be accessible Electronic Information and/or Electronic Records with contents against propriety”.

“ Any Person who knowingly and without authority distributes and/or transmits and/or causes to be accessible Electronic Information and/or Electronic Records with contents of affronts and/or defamation”.

“Any Person who knowingly and without authority distributes and/or transmits and/or causes to be accessible Electronic Information and/or Electronic Records with contents of extortion and/or threats.”.

if it violates these articles it will be punished according to article 45 paragraph 1 which reads:

“Any Person who satisfies the elements as intended by Article 27 section (1), section (2), section (3), or section (4) shall be text to imprisonment not exceeding 6 (six) years and/or a fine not exceeding Rp1,000,000,000 (one billion rupiah).”.

In terms of religion, especially Islam, cyberbully acts are included in despicable acts. Islam strictly prohibits and strongly discourages other people's degrading behavior. This is as explained in the word of Allah Almighty in Surah Al-Hujurat verse 11:

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا يَسْخَرْ قَوْمٌ مِنْ قَوْمٍ عَسَىٰ أَنْ يَكُونُوا خَيْرًا مِنْهُمْ وَلَا نِسَاءٌ مِنْ نِسَاءٍ عَسَىٰ أَنْ يَكُنَّ خَيْرًا مِنْهُنَّ وَلَا تَلْمِزُوا أَنْفُسَكُمْ وَلَا تَنَابَرُوا بِالْألقَابِ يَنُسُ الإِسْمُ الفُسُوقُ بَعْدَ الإِيمَانِ وَمَنْ لَمْ يَتُبْ فَأُولَئِكَ هُمُ الظَّالِمُونَ

Meaning: "O you who have believed, let not a people ridicule [another] people; perhaps they may be better than them; nor let women ridicule [other] women; perhaps they may be better than them. And do not insult one another and do not call each other by [offensive] nicknames. Wretched is the name of disobedience after [one's] faith. And whoever does not repent - then it is those who are the wrongdoers. " (Q.S. Al-Hujurat: 11),

Nixon (2014) revealed that the cyberbully effect has been extensively explored in the field of adolescent mental health problems. In previous research explained that there are relationship between involvement with cyberbully and the tendency of adolescents to internalize problems (for example, development of negative affective disorders, loneliness, anxiety, depression, suicidal ideas, and somatic symptoms). Therefore, As way of decrease number of cyberbully crime, the research propose a study entitled ‘Detection of Textual Cyberbully by using the SVM Algorithm. The final result form of our research is an API (Application Program Interface) that will be freely used by all social media in Indonesia, so that it can become a standard so that texts detected as social media will be handled. And with this, it is expected to reduce the culture of Indonesian society to do cyberbully crimes.

1.2 Research Question

1. How accurate is the Support Vector Machine (SVM) algorithm to detect textual cyberbully?
2. How fast is the Support Vector Machine (SVM) algorithm to detect textual cyberbully?
3. How reliable is the Support Vector Machine (SVM) algorithm to detect textual cyberbully?

1.3 Research Objectives

1. To measure the accuracy of Support Vector Machine (SVM) algorithm to detect textual cyberbully.
2. To measure the speed of Support Vector Machine (SVM) algorithm to detect textual cyberbully.

3. To measure the reliability of Support Vector Machine (SVM) algorithm to detect textual cyberbully.

1.4 Research Scope

1. Research only focuses on the detection of textual cyberbully in Bahasa Indonesia.
2. Data is collected from Instagram posts.
3. Research only focuses on the word content. While other aspects such as age, environment, etc are ignored.
4. The text will be divided into two groups, namely the cyberbully text and non cyberbully text.

1.5 Research Benefit

The output of this research will be valuable for the developer of social media, that they can use free API to minimize function in their program to reduce cyberbully. While the benefit for the wider community is that it can reduce the number of cyberbully crimes.

CHAPTER II

LITERATUR REVIEW

2.1 Detection of Cyberbully

Patchin & Hinduja (2012) defines cyberbully as an intentional activity carried out repeatedly which poses a danger that is carried out through the use of computers, cellphones, and other electronic devices. Some examples of cyberbullyin is sending insulting, demean, embrassing, and threatening messages, deceiving someone to reveal personal or embarrassing information and sending it to others, make a fake profiles on social media for fun other people. There are several forms of cyberbully as stated by Willard (2007) :

a. Flaming (spread dispute), which is when an initial dispute occurs between two or more people (on a small scale) and then spreads so that it involves many people (on a large scale) so that become a big noise and problem.

b. Harrasment, which is an attempt by someone to harass another person by sending various forms of messages both in writing and images that are hurtful, insulting, shameful and threatening.

c. Denigration (slander), which is a person's effort to spread false news that aims to damage the reputation of others.

d. Impersonation (imitation), namely the effort of someone pretending to be someone else and seeking a third party tell things that are confidential.

e. Outing and trickery (fraud), namely the effort of someone who pretends to be someone else and spreads false news or the secrets of the other person or third party.

f. Exclusion, which is an attempt to isolate or exclude someone from joining a group or community for discriminatory reasons.

g. Cyber-stalking, namely the effort of someone stalking or following someone else in cyberspace and causing interference to the other person.

Dinakar (2011), explained that text cyberbully has pattern as follows:

- a. Sexuality: Negative comments involving attacks on sexual minorities and sexist attacks on women.
- b. Race and Culture: Attacks bordering on racial minorities (e.g. African-American, Hispanic and Asian) and cultures (e.g. Jewish, Catholic and Asian traditions) including unacceptable descriptions pertaining to race and stereotypical mocking of cultural traditions.
- c. Intelligence: Comments attacking the intelligence and mental capacities of an individual.

Pang (2013) shows that there is a correlation between pronoun and bully word on labeling cyberbully text. And Hariani (2017) that shows that Indonesian society often uses the animal word to bully someone.

Here are some studies on cyberbully:

Nahar and Zhang (2013) proposed a session-based framework for automatically detect cyberbully from the huge amount of unlabelled streaming text

and in real world situation. They investigated the one-class ensemble learning method for cyberbully detection to tackle the real world situation. they train their system using a small set of positive trainings only, where the system automatically extracts reliable negative and more strong positive samples for training from the huge amount of unlabelled data. Nevertheless, the results of their research is indicated that a feasible means to learn cyberbully instances, effectively and automatically, from unlabelled text streams from SNs.

Singhal and Bansal (2013) giving a survey on cyberbully according physical appearance, race and ethnicity ,sexuality and sexual identity. They design and implement a social networking web site by which researcher can simulate bullying, and prevent the users to get bullied. they provide a case study and the performance study of our system. They represent a survey on the some scenario of cyberbully and various methods available for the detect and prevent cyber harassment. Their concept is upload text written by any user is first analyzed and after that, they estimate the roles of user and then they provide help as required by the user using data mining techniques. User Identity for registration on their site is only one, they provide an identity proof for registering on their site. Result of this research is not shown clearly.

Potha and Mouragodakis (2014) proposed utilizing a dataset of real world conversations in which each predator question is manually annotated in terms of severity using a numeric label. They approach the issue as a sequential data modelling approach and use a Singular Value Decomposition representation for formulate predator's questions. they exploit the whole question set and model it as a signal, whose magnitude depends on the degree of bullying content. Use feature

weighting and dimensionality reduction techniques for each signal is straightforwardly parsed by a neural network that forecasts. the result of their research is that it confront the task as a time series modeling approach, in which previous states portray important information on the knowledge of a future state.

Hon and Varathan (2015) tried to solve the problem of cyberbully by conducting researc with data that is different from previous studies, they prefer data from twitter. The method used is the method of NGOs. The purpose of their research is to monitor the social media community, generally on Twitter victims and get data so that they can prevent subsequent cyberbully. The strength of their research is identifying the intentivity and information of cyberbully actors, also using location data, and reports from Twitter users. But the results of this study are less satisfactory with the results of low accuracy.

Van *et al.* (2015) study to detect cyberbully texts contained in texts automatically with qualitative classification of analysis. The purpose of their research is to provide information or data about cyberbully that is presented based on data that has been obtained from their research. The strength of their research is to use many methods to detect cyberbully. Among them are SVM and naive bayes. The shortcomings of this study are the only results of the accuracy of this study which only reached 53.82%.

Algaradi *et al.* (2016) wanted to perfect the research conducted by Varathan (2015) which detected online communication, namely twitter using libSVM, NB, random forest. Their purpose of this research is to provide a possible solution to real time online communication. The strength of this research is being able to handle

data using the dataset they have made. The weaknesses are still the same as previous studies, it is the weak level of result accuracy.

Savanaraj *et al.* (2016) tried to solve the problem that Algaradi *et al.* (2016) wanted to solve use are sintatic and semantic techniques that use naive bayes and random forest for their classification. The purpose of conducting this research is that they try to reduce the victims of rounding up on social media, the same goal as all researchers who conduct research in the field. The strength of their research is the integration between detection and rumor in one application. While the shortcomings of their research are that they can only be implemented in one application and that accuracy is low.

Zhong *et al.* (2016) study cyberbully detection using data that is slightly different from previous research. The datas is images. The method they use is using SIFT, CH and GIST. The purpose of their research is to reduce rounding problems or in terms of images inputted by users on Instagram social media. They also process caption data entered by the user. The strength of their research is in the form of processed data that is very different from the data that has been studied in previous studies. But just like the research that has been done so far, the results of the research are not satisfactory, with proven from low accuracy.

Haidar *et al.* (2017) research to solve cyberbully problems with built with the aim of completing from 2 languages namely English and Arabic. The method that they use is a hybrid traning model such as a combination of distance function, NB and SVM. the initial goal of their research was actually to detect texts, when the text uses Arabic. The strength of their research is that they focus on detecting texts

in their language, namely Arabic. But they are able to provide data accuracy that is quite high at 94%. The weakness of their system is that the value of accuracy is only calculated based on the success of the language, but not on the new data in Arabic.

2.2 Support Vector Machine

Support Vector Machine (SVM) was developed by Boser, Guyon, Vapnik, and first was presented in 1992 at the Annual Workshop on Computational Learning Theory. The basic concept of SVM is actually a harmonious combination of theories computing that has existed for decades before, like the hyperplane margin (Duda & Hart 1973, Cover 1965, Vapnik 1964, etc.), the kernel was introduced by Aronszajn in 1950, and so too with supporting concepts other. But until 1992, there had never been an attempt to assemble that components. Unlike the neural network strategy that tries to find a separating hyperplane between classes, SVM tries to find the best hyperplane in input space.

It is assumed that the two classes can be completely separated by the hyperplane in D-dimensional feature space. The Hyperplane is defined as follows:.

$$w * xi + b = 0 \quad (2.1)$$

The x_i data belonging to the negative class are those who fulfill the inequality the following:

$$w \cdot x_i + b \leq -1 \quad (2.2)$$

The x_i data belonging to the positive class are those who fulfil inequality:

$$w \cdot xi + b \geq 1 \quad (2.3)$$

Optimal margins are calculated by maximizing the distance between the hyperplane and pattern closest. This distance is defined as $2/|w|$ ($|w|$ is a norm of weight vector w).

Some reason using SVM for detection is :

1. Generalization Generalization is defined as the ability of a method (SVM, neural network, etc.) to classify a pattern, which does not include data used in the learning phase of that method. Vapnik explained that the generalization error is influenced by two factors: an error in the training set, and one more factor that is influenced by the dimensions of VC (Vapnik-Chervokinensis). Learning strategies on neural networks and generally learning machine methods are focused on efforts to minimize errors in training-sets. This strategy is called Empirical Risk Minimization (ERM). As for SVM, besides minimizing errors in training-sets, it also minimizes the second factor. This strategy is called Structural Risk Minimization (SRM), and in SVM it is realized by choosing a hyperplane with the largest margin. Various empirical studies show that the SRM approach to SVM gives smaller generalization errors than those obtained from the ERM strategy on the neural network and other methods.

2. Curse of dimensionality Curse of dimensionality is defined as a problem faced by a pattern recognition method in estimating parameters (for example the number of hidden neurons in a neural network, stopping criteria in the learning process etc.) due to the relatively small number of data samples compared to the dimensional

vector space data . The higher the dimensions of the vector information space that is processed, the consequences of the need for the amount of data in the learning process. In reality it often happens, the data processed is limited in number, and to collect more data is not possible because of cost constraints and technical difficulties. In these conditions, if the method is "forced" to work on data that is relatively small in number compared to its dimensions, it will make the process of estimating the parameter method very difficult. Curse of dimensionality is often experienced in applications in the field of biomedical engineering, because usually available biological data is very limited, and its provision requires high costs. Vapnik proves that the level of generalization obtained by SVM is not influenced by the dimensions of vector input [3]. This is the reason why SVM is one of the right methods used to solve high-dimensional problems, within the limitations of existing data samples.

3. Theoretical basis As a statistical-based method, SVM has a theoretical basis that can be analyzed clearly, and is not black box.

4. The SVM feasibility can be implemented relatively easily, because the process of determining support vectors can be formulated in the QP problem. Thus if have a library to solve QP problems, SVM can be implemented easily by itself. In addition, it can be solved by sequential methods as explained earlier.

Here are some studies on Support Vector Machine :

Zheng *et al.* (2012) conducted research for fruit sorting according to the automatic maturity level of mangoes using LS-SVM as a function of fractal analysis and CIELAB parameters. The aim of the study was to evaluate and compare the

performance of LS-SVM to be grouped based on fractal analysis. The parameters used in this study are color, FD, dimensions, and dimensions of dilation. The conclusion of this study shows that the use of LS SVM method can be used to detect fruit maturity because it has an accuracy value of 88.89%. The lack of this research is that the use of research results for different fruits has different values.

Nayak *et al.* (2015) conducted a study to detect brain abnormalities in Magnetic Resonance Imaging (MSI) using Multirelution analysis by utilizing the SVM Least Square algorithm (SVM LS). By utilizing the multiresolution technique called DWT, they perform feature extraction in the brain MRImaging. Then the data is calculated to reach level 3 coefficients which then they do the feature matrix generate. After the feature matrix is obtained, they normalize the features, and reduce them using KPCA, then the final step, ie the images obtained are classified using the SVM algorithm. The results of this study are that they claim that the SVM algorithm is very efficient at detecting brain abnormalities, as evidenced by the high accuracy of up to 100% with RBF, 94.44% for linear, and 98.89% for polynomials.

Susiloewati *et al.* (2015) tried to implement SVM to classify traffic congestion on Twitter. In their research, SVM was tested to find out the performance in classifying data in the form of text which then the classification data was transformed into the form of Google Map visualization. Their research model is to use several scenarios, starting from the scenario in the dataset, to the composition of training data and testing data. The results of their research show that in some scenarios, or more precisely the first scenario they get an accuracy of 100 percent, while in the second scenario, the greatest accuracy is only indicated by the value of 93.73%. Their research states that there are many mistakes in their research design

such as the inappropriate training data and they have difficulty determining the right keywords to be able to screen data until it can be processed correctly. They also revealed that the higher the amount of data used in the training process, the higher the level of accuracy.

Munawaroh R. (2016) diagnosed hepatitis C disease using SVM Algorithm, this study uses testing data with two s by using 100 positive data training data and 100 negative data. The features used in this study use as many as 8 features, which means 8 dimensions, so that you need to use the , and the used in this study is a linear and RBF . The results of the study using linear functions get the correct percentage results of 68-83% and RBF functions 70-96%. The conclusion of this study is that the SVM method can be used to diagnose hepatitis with a high degree of accuracy and RBF functions have a level of accuracy tend to be higher than linear functions.

Zidi *et al.* (2017) try to use SVM to detect errors on the WIFI network sensor. In this study, the SVM method was applied in order to classify the data received from the sensor, and detect sensor errors. Their research model is that they prepare data and look for the most optimal value on SVM parameter values, they try to use the matlab box repeatedly until the optimal value is found with alpha 0.5 and C equal to 1. Then after getting the maximum value, they compare with the technique - different techniques and methods to test the abilities of their research results. The results of using SVM to detect errors on the WIFI network sensor have a fairly high and stable accuracy above.

Purnamasari *et al.* (2018) Tried to do research distinguish texts using 220 data. They got the results obtained from all SVM sequential training parameters, with $\text{iterMax} = 20$, $\lambda = 0.5$, $\gamma = 0.001$, $\varepsilon = 0.000001$, and $C = 1$. The accuracy results obtained were 75% accuracy, 70.27% precision, recall 86.66% and f-measure 77.61%. And on the results of testing the information gain feature threshold, the best results obtained are at the threshold value of 90%, with an accuracy value is 76.66%, precision is 72.22%, recall is 86.66% and f-measure is 78.78%. The disadvantages of being selected feature cause overfitting.



CHAPTER III

SYSTEM DESIGN AND IMPLEMENTATION

3.1 Data Collection and Analysis

The data collection and analysis is illustrated in Figure 3.2. the process as elaborated with several steps. First step in data collection is login social media and then search and choose user which is frequently occurs cyberbully. In this research, the social media used is Instagram. the reason for using Instagram is because:

1. According to Ditch the label (2017) Instagram is the social media that has the most cyberbullying, which is as much as 42%.
2. Jakpat (2018), shows that Instagram is most interesting social media by teenagers and young users in Indonesia, which is 61.9%.

Second step in data collection is crawl text comment from selected account. In this research is @lucintaluna and @ahmadhaniprast account had been chosen because their account became trending topic at beginning year of 2018. Then, third process is to choose several texts from comment list in data. Fourth process is data analysis, where main form of data analysis is data validation. And the last process is to save data.

To data analysis, the research spread a questionnaire containing 153 comment text for respondents to choose. The criteria for respondents are college student and have Instagram account. Then the data obtained is separated according to the majority. Data is separated into 2 labels, namely the label cyberbullying and non cyberbullying. We will use the data as a reference for training and testing. 122 for training and 31 for testing. After that, text and valid data save in database with

JSON form, where text save in text column and result type from questioner save as target column.

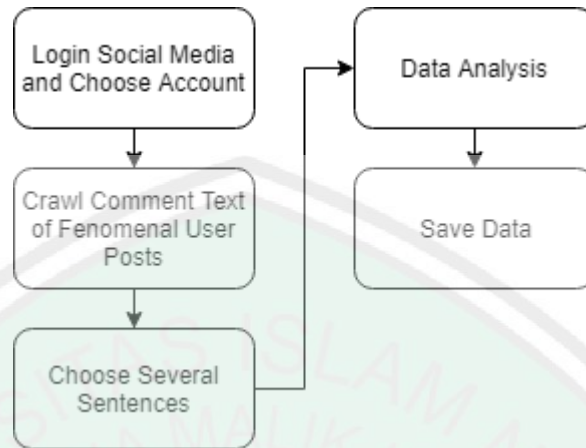


Figure 3. 1 Data Collection Flow

3.2 System Design and Implementation

The design of this detection system is shown in Figure 3.5. The process can be elaborated as follows: Input data. Next is data is processed on preprocessing state. After that, the process is feature extraction process, the main purpose of this process is to get the value of the predetermined feature. Then, the system conducts training to get the hyperplane, this hyperplane will determine which type of the text, as cyberbully text or non cyberbully text. After that, system save the hyperplane in database. Next process is system conduct a testing process using a hyperplane which got in training process. Output as result of this testing is values between -1 or 1 where -1 as noncyberbully text and 1 as cyberbully text.

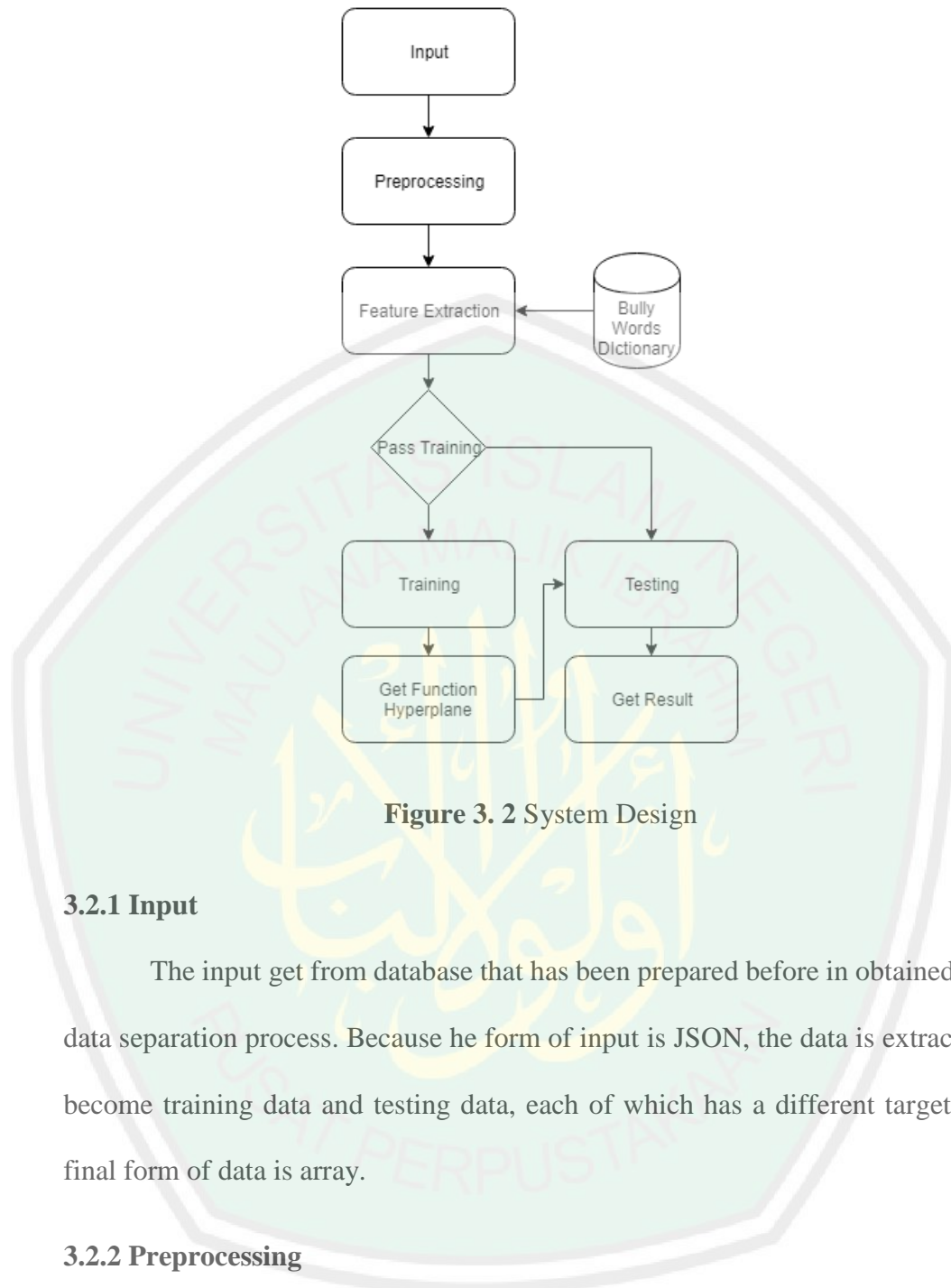


Figure 3. 2 System Design

3.2.1 Input

The input get from database that has been prepared before in obtained from data separation process. Because he form of input is JSON, the data is extracted to become training data and testing data, each of which has a different target. And final form of data is array.

3.2.2 Preprocessing

According Hariani (2017), The process of data mining can be carried out if the data into focus before. The process like cleaning, throw out the duplication of data. Our preprocessing process is shown in Figure 3.4, is the same as the general design with the aim that the text will be ready to be used in the next process, in other words to fulfill the text requirements it can be processed properly in the next

process. process steps can be described as follows: The preprocessing process start with inputting the text will be detect, both training data and testing data both pass the preprocessing process. Next process is to do the cleaning process, where the illustrate of cleaning process can be shown in Figure 3.3. Next process is texts that are inputted into the process are deleted by special characters such as : , ? = and other. Next process is make all letters of text into lowercase letters so text can be processed more precisely.

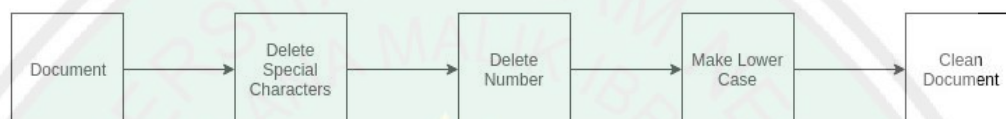


Figure 3. 3 Text Cleaning Design

After the cleaning process is completed, the next process is to remove stopword, where this remove of stopword can be done to avoid error in weighting a feature that will affect error detection. Some examples of words that are discarded are conjunctions like and with and others. after the stopword process, researcher did the second cleaning process, researcher did it because in the stopword process researcher built sometimes left an accidental error, then as an alternative researcher did the second cleaning process. For stopword process, the dictionary we use is using the php library provided by Satrawi which can be accessed on the link: <https://github.com/sastrawi/sastrawi>.

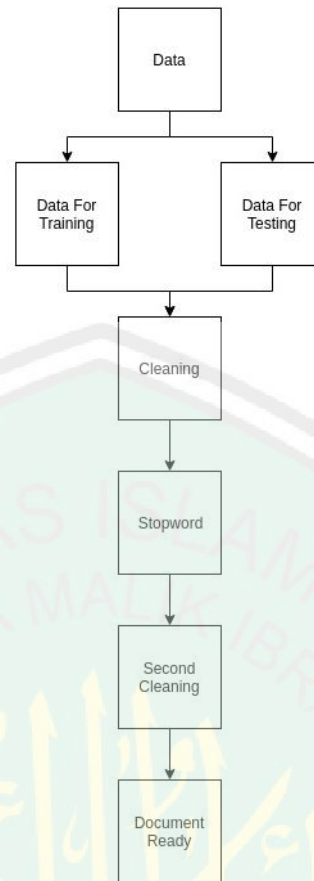


Figure 3. 4 Preprocessing Design

3.2.3 Feature Extraction

Feature extraction aim to find key of significance in the text depending on their intrinsic characteristics. feature extraction is obtained by looking for things that distinguish between 2 or more things, such as text content, text length, etc. Feature extraction in this research is a process for extracting features can be explained by Figure 3.5, the first process of texts that have gone through the preprocessing process and then looking for features using dataset prepared and analyzed in our second research process.

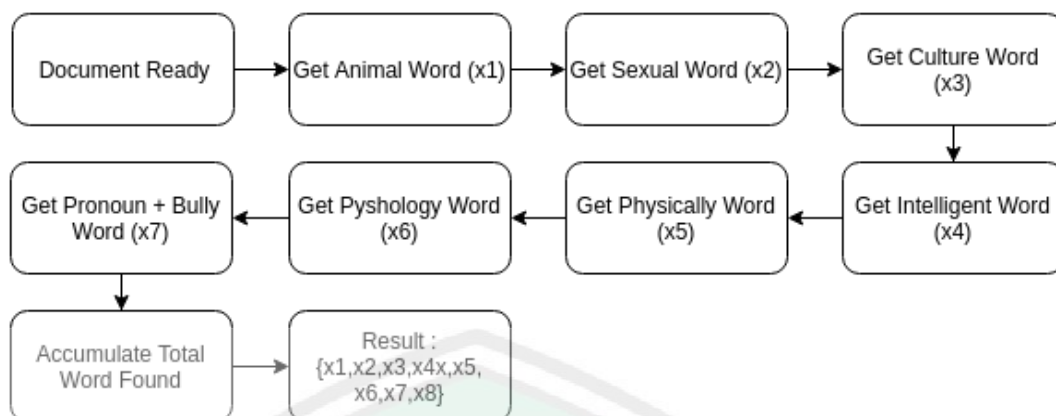


Figure 3. 5 Feature Extraction Design

The features used in this study basically use the existence of bully words in a text that is processed in such a way, while the list of features that will be used is shown in table 3.1.

Table 3. 1 List Feature

No	Features
1	Amount of Animal Word
2	Amount of Sexuality Word
3	Amount of Race/Culture Word
4	Amount of Intelligent Word
5	Amount of Physically Word
6	Amount of Psychology Word
7	Amount of pronoun + bully word / Amount of bully word + pronoun

8	Total bully / Total Texts
---	---------------------------

3.2.4 Training

In this research, the SVM model used is a linear model, a linear model is used because the research model separates cyberbully text from non-cyberbully text in a linear manner which only uses the existence of bully words themselves. the pattern of data is very simple, that is, the data contained in the dictionary of cyberbullying will be suspected of being cyberbullying, and which is not necessarily cyberbullying. then, between texts that contain bad words will be distinguished only by the number or type. and some types will have different weights which will determine whether the text is bully or not. to distinguish these texts, the linear function is the best, because the weighting process carried out previously has determined the final function, without having to give a different value to the operation that is in a function other than linear.

Which is where the existence is processed in such a way as to obtain some features that can be used for research. The formula used is the simplest formula or the basic concept of svm. The formula is :

$$F(y) = a * X + b \quad (3.1)$$

Where,

a = the weight value obtained from the training process then to get the best margin then the formula used is $2/|W|$

X = the sum of each feature that has been paired with the parameter value, X is f(y) result from previous training.

b = bias

by using the hyperplane if $f(y) < 0$ then, the results obtained are texts not including cyberbully, whereas if $f(y) > 0$ then, the results obtained by the positive text are cyberbully.

The training process for obtaining the parameter value of the feature has explained previous, then the training process can be explained with the following flowchart at Figure 3.6.

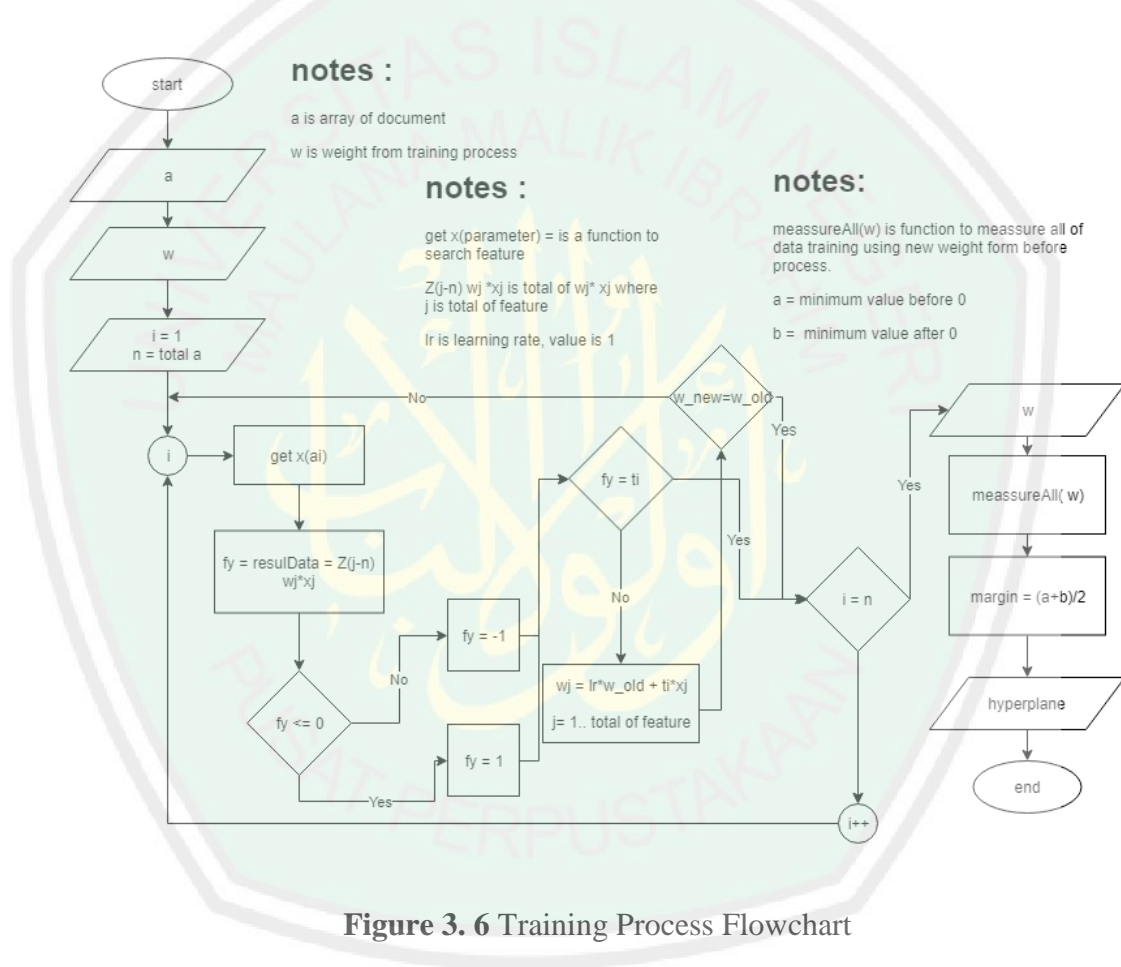


Figure 3. 6 Training Process Flowchart

The training process was begins with the feature value of all feature that has been obtained from the feature extraction process. Then the parameter weight values are initiated first, with values:

$$w \text{ first} = 1/\text{total of feature} \tag{3.2}$$

in this study there were 7 features then value from w , for first compute is $1/\text{feature total}$. Then test the system to find the exact parameter value in the way:

1. The function created is

$$f(y) = \sum_{i=1}^{\text{total feature}} w_i * x_i \quad (3.3)$$

Where :

w_i : weight from each of feature

x_i : parameter value from each of feature

2. Then the $f(y)$ results are searched, if the $f(y) < 0$ then it is initiated as -1, if instead it is initiated as 1

3. If the value of $f(y)$ is equal to the target then the value of w does not need to be updated

4. If the value of $f(y)$ is not the same as the target, the value of w is updated, using the formula:

$$w_{\text{new}} = w_{\text{old}} + \text{target} * x_s \quad (3.4)$$

5. The iteration continues until w_{new} is not change.

6. After that, system completely measure all of data training, then get a minimal value before 0 and after 0. Then sum both of value and divide by 2 then use that value as a margin. The formula like :

$$c = \frac{(a + b)}{2} \quad (3.5)$$

Where,

margin is as a limit from hyperplane,

a is value minimum before zero,

b is value minimum after zero.

The result of this training is the value of each parameter (w). the value of each parameter will various. With the parameter, hyperplane function can be created.

And hyperplane, can be written as:

$$y \left(\sum_{i=1}^{\text{total feature}} w_i * x_i \right) \geq c \quad (3.6)$$

Where,

$y = 1$, if $x > c$

and $y = -1$ if $x < c$

3.2.5 Testing

The testing process is the same as the training process only, this process has used parameter values for features obtained from the training process. All of step shown at Figure 3.7.

The working conditions of this process are:

1. The text you want to test is entered
2. Then extract the feature

3. The value of feature extraction is processed using the formula 3.6

5. If the result of $f(y) < \text{margin}$ then the text is classified as a non-bully text, If otherwise, texts are classified as bully texts.

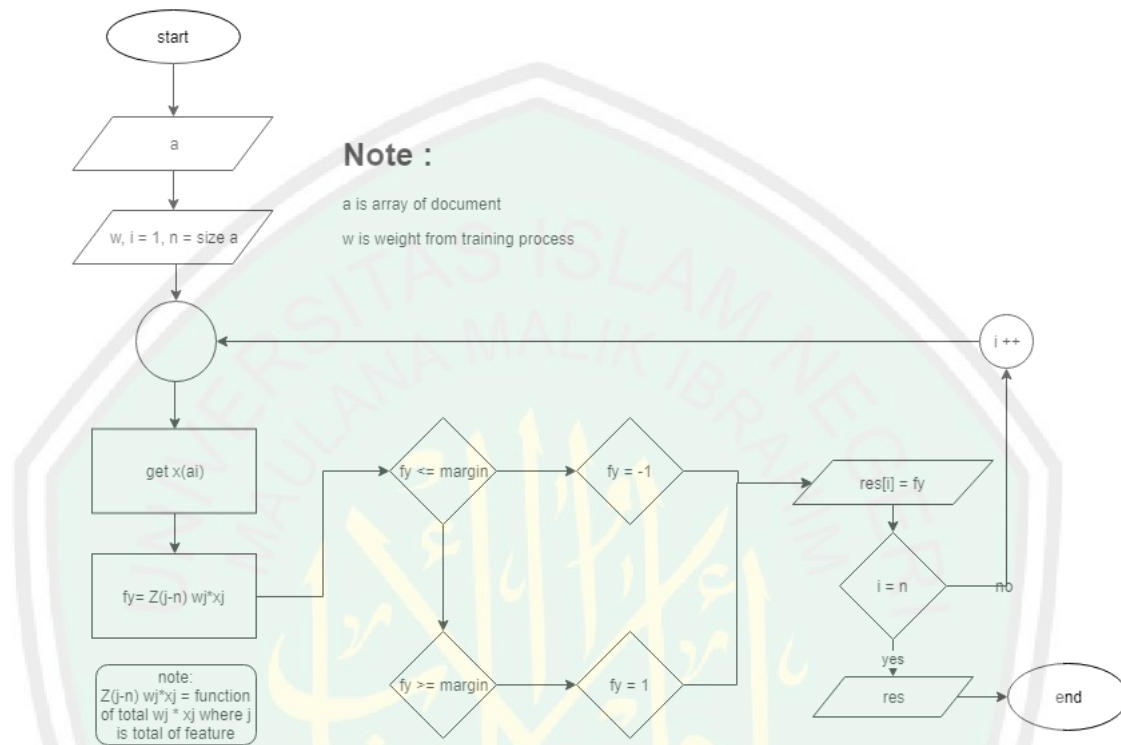


Figure 3. 7 Testing Process Flowchart

The method proposed previously was implemented in the form of an API written in PHP version 5.6.20 using the slim library using a visual studio code editor. Then this API will be accessible using a website built using a library code igniter. The application was built using a laptop equipped with Microsoft Windows 10 OS with Intel Celeron B815 processor and 4 GB RAM. Implementation is finished by writing down the program functions described in :

3.2.6 Data Crawl

To crawl Instagram data using the python program provided by Timgrossman, sources can be accessed on the following website: <https://github.com/timgrossmann/instagram-profilecrawl>. The data that is crawled is from lucintaluna, ahmaddhaniprast accounts.

3.2.7 Database Design

Database designed using MySQL database. Database consist of several tables. Each table stores data needed for all detection process and for displaying the system, including storing user data. Database design for this system illustrated in Figure 3.8.

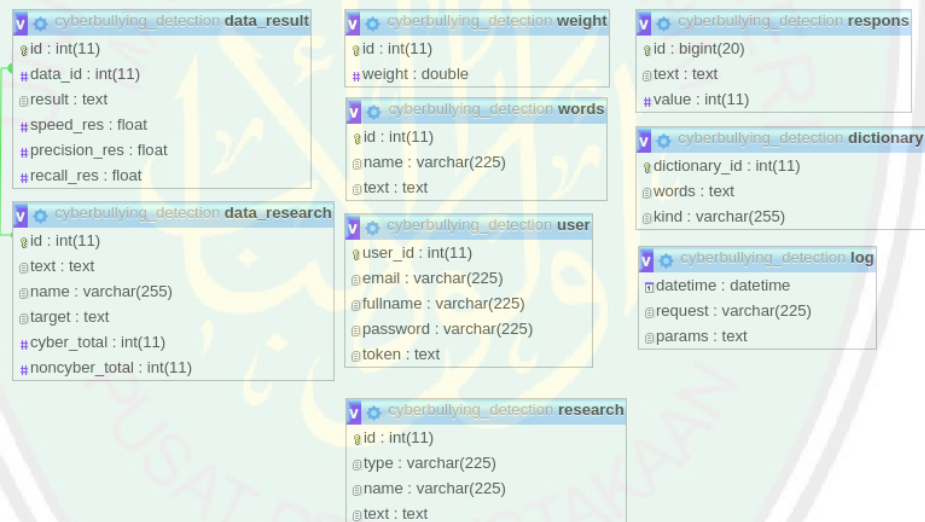


Figure 3. 8 Database Design

Based on figure 3.8 can be explained as follows:

1. Table `cyberbullying_detection_data_research` function was designed to store data used in research both of training data and testing data

2. Table `cyberbullying_detection_data_result` function is to store data result after processed.
3. Table `cyberbullying_detection_weight` function is to store parameter value after training process.
4. Table `cyberbullying_detection_words` function is to store words used in process like stopwords.
5. Table `cyberbullying_detection_user` function is to store user data in register process.
6. Table `cyberbullying_detection_respons` function is to store respons user about system.
7. Table `cyberbullying_detection_dictionary` function is to store dictionary data.
8. Table `cyberbullying_detection_log` function is to store log data in trial process.
9. Table `cyberbullying_detection_research` function is to store backup data for process training and system.

3.2.8 Input Data Implementation

The main process in this program is two, is training process and training process, each of which will have different data. Data training and data testing are stored in table 3.2.

Table 3. 2 Database data

name column	Type	Length
id	Int	11
text	Text	
name	Varchar	255
target	Text	
cyber_total	Int	11
noncyber_total	Int	11

Explanation of table 3.2 can be described as follows, the id is to store the id of the text, then type to store the type of text namely training or testing, name to fill 2 types, namely texts or targets, while the text contains a collection of texts that will be used in research.

Then retrieve data using the PDO interface to communicate with MySQL database. For data communication code and data calling code shown in Figure 3.12 and Figure 3.13.

```
public function get( ) {
    $user_secret = $this->db->prepare('
        SELECT name,text FROM research
    ');
    $user_secret->execute( );
    return $user_secret->fetchAll(PDO::FETCH_ASSOC);}

```

Figure 3. 9 Get Data from Database

```
// get data training and data target from db
$data = $this->research_model->get();
$data = $this->_loopWord($data);

```

Figure 3. 10 Communication Data

3.2.9 Generating Dictionary

Dictionary data is stored in mysql database with table format as shown in table 3.3:

Table 3. 3 Dictionary Table

name column	Type	Length
id	Int	11
word	Text	
kind	Varchar	

The explanation of table 3.3 is as follows: dictionary_id is used to put the id of each feature stored in db. Whereas words are a column to store a collection of words, whereas kind is the name of the feature.

Implementation of generating dictionary by writing a script for calling data and preparing it in an array to be used in the next process. The code script generating dictionary can be seen in Figure 3.10.

```

public function _generatingDictionary(){
    // get list word from db
    $dictionary = $this->dictionary_model->get();
    // trim data from db (clean enter and spaces)
    $this->dictionary = $this->_loopWord( $dictionary ); }

// Function Loop Word
// Preprocessing text from db
public function _loopWord($data, $index = 0 ){
    // explode by enter
    $data[$index]['text'] = explode(PHP_EOL,
    $data[$index]['text']);
    // trim(delete spaces)
    $data[$index]['text'] = $this->_trim($data[$index]['text']);
    if( $index == sizeof($data) - 1 ) return $data;
    return $data = $this->_loopWord( $data, $index + 1 );}

```

Figure 3. 11 Generating Dictionary Script

3.2.10 Preprocessing Implementation

After the data is obtained and ready to be used, the next process is the preprocessing process. Preprocessing implementation consists of several processes as described in the previous discussion, namely cleaning, stopword removal, and second cleaning. Examples of preprocessing data before processing and after processing can be seen in Figure 3.11 and Figure 3.12.

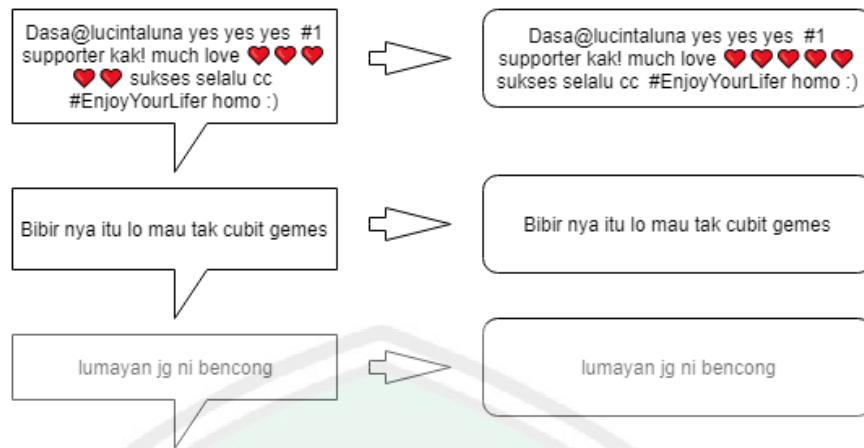


Figure 3.12 Data Before Preprocessing

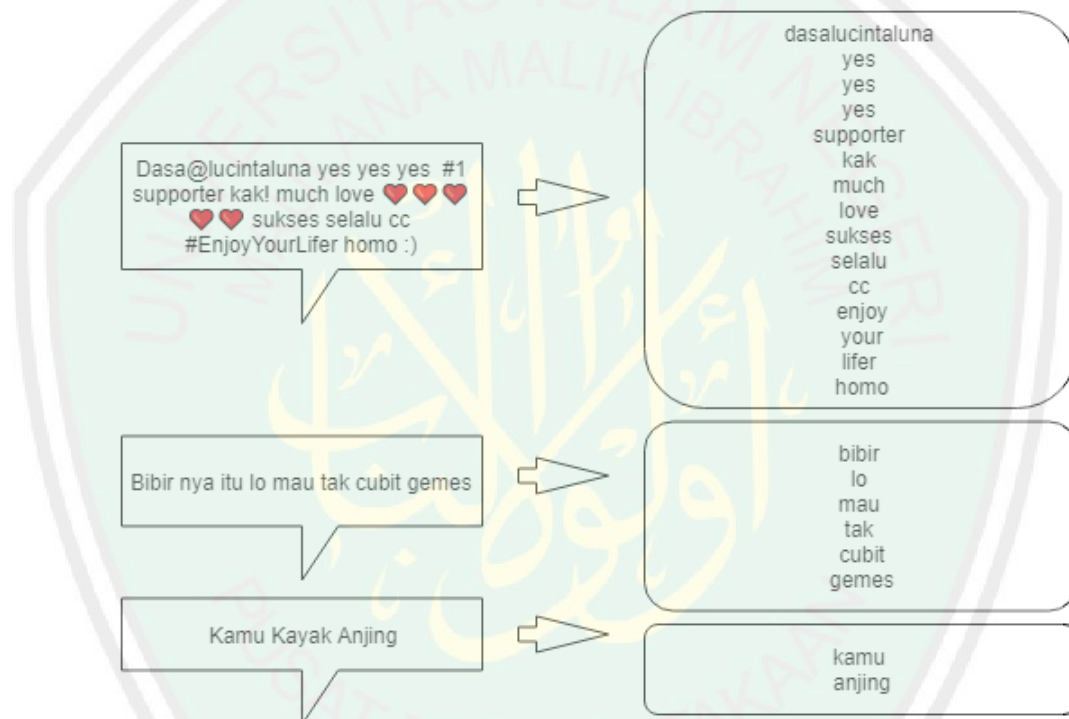


Figure 3.13 Data After Preprocessing

For implementation, the code is separated into several functions so that it can be used multiple times and the code can be seen in Figure 3.13.


```

public function _preprocessing( $data ){
    // cleaning
    $data = $this->_cleaning( $data );
    // stopword
    $data = $this->_stopword( $data );
    // second cleaning
    $data = $this->_secondCleaning( $data );
    return $data;
}

```

Figure 3. 14 Preprocessing Script

1. Cleaning

This process is a process which removes numbers, characters, or spaces that are not needed in the detection process. The example of cleaning process can be illustrated in Figure 3.14.

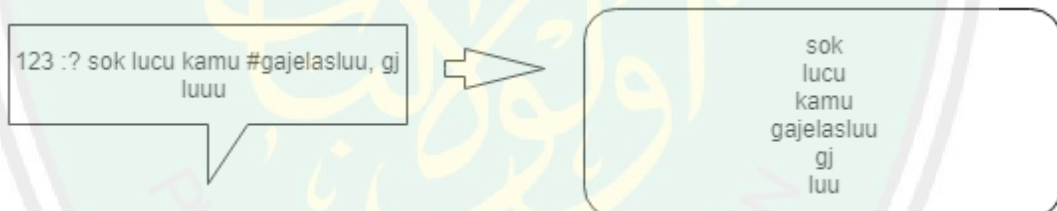


Figure 3. 15 Example Cleaning in Text

Implementation of cleaning done with write script code of cleaning data. And script code can be seen in Figure 3.15.

```
function _cleaning( $data ){
    $data = str_replace("/r","",$data);
    $data = preg_replace('/^[^A-Za-z0-9\ ]/', '', $data);
    // delete spesial karakter kecuali spasi
    $data = preg_replace('/[0-9]+/', '', $data);
    $data = strtolower($data);
    return $data;
}
```

Figure 3. 16 Cleaning Script Code

2. Stopword

The stopword process is a process where you remove unnecessary words or words that can interfere with system performance. The stopword process can be illustrated in Figure 3.16.

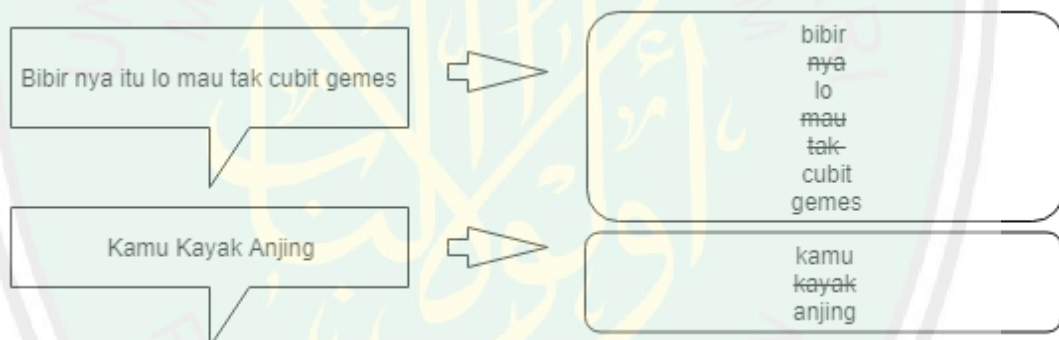


Figure 3. 17 Example Stopword in Process

Main of stopword process is delete some word if not useless for program. Stopword list saved in database with name stopword, and data saved in text type data. The script code of process stopword can be seen in Figure 3.17.

```
function _stopword( $data ){
    //remove stopword
    $data =
    preg_replace('/\b(' . implode('|', $this->stopword_list) . ')\b/', ' ', $data);
    return $data; }

```

Figure 3. 18 Stopword Script Code

3. Second Cleaning

This process is the final cleaning process of previous processes when something is missing like space, or blank word. The example of second clening illustrated in Figure 3.18.

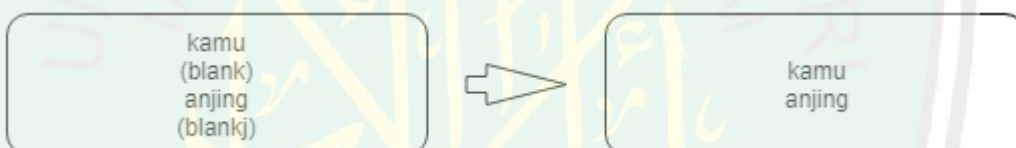


Figure 3. 19 Example Second Cleaning in Text

Main purpose of second cleaning is delete some useless character or trash like blank space form process before. Script code of second cleaning can be seen in Figure 3.19.

```
function _secondCleaning ( $data ){
    $data = preg_replace('/\s\s+/', ' ', $data);
    $data = trim($data);
    return $data; }

```

Figure 3. 20 Second Cleaning Script Code

3.2.11 Get Feature of Text (Extraction of Feature)

The value extraction process from features that have been determined in the previous discussion, namely the total value of the word animal, sexuality, race, intelligent, psychology, physically, and pronoun. The feature extraction process is carried out both in the training process and in the testing process. In the process of training and testing that involves a lot of it, the extraction process uses a looping function, while for single text testing, it only takes one time through the feature extraction process. Examples of feature extraction results from texts can be seen in Figure 3.20.

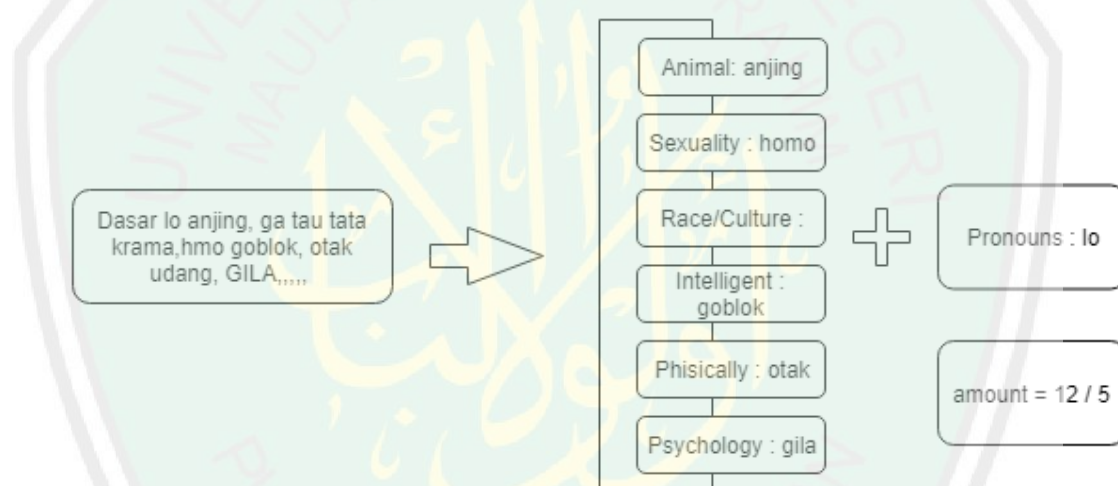


Figure 3. 21 Example Get Feature of Text

For script implementation, make several extraction processes to avoid duplicate code. In some functions, use recursive loops to loop texts, or features. The code script can be seen in Figure 3.21.

```

public function _getFeature( $explodeSentence, $featureObtained =
array(), $index = 0 ){

    $featureObtained =
$this->_searchFeatureInDB( $explodeSentence[$index],
$featureObtained);

    // SPECIAL FOR PRONOUN
    // If no bully word, delete pronoun was found
    if( $index == 6 ){
        if( sizeof($featureObtained) == 1 ){
            if (array_key_exists(6, $featureObtained)) {
                $featureObtained = null;
            }else return;
        }else return;}
    if($index == sizeof( $explodeSentence ) - 1) {
        return $featureObtained;
    }
    return $this->_getFeature( $explodeSentence, $featureObtained,
$index + 1 );}

function _searchFeatureInDB( $word, $featureObtained, $index = 0 ){
    $obtained = array_search( $word,
$this->dictionary[$index]['text' ] );
    if( $obtained !== false ){
        $this->totalBully = $this->totalBully + 1;
        if(!isset($featureObtained[$index])){
            $featureObtained[$index] = array();
        }
        array_push($featureObtained[$index], $word);
        return $featureObtained;}
    if($index == (sizeof( $this->featureInDB ) - 1)){
        return $featureObtained;}
    return $this->_searchFeatureInDB( $word, $featureObtained,
$index + 1 );}

```

Figure 3. 22 Get Feature Script Code

3.2.12 Training Implementation

Training implementation is done by writing a training script, where the training script returns the weight value that will be used in the testing and detection process. In the training process, it involves many looping functions because there are several processes that need to be repeated to achieve the desired values as described in the research design chapter. The final result of the training process is that the hyperplane function is used to separate texts including cyberbully or not cyberbully.

The training process is structured by several processes that have been discussed previously, adding several processes. The training script consists of several functions according to the one described in the flowchart training in the research design. The following are some of the processes in the training process:

1. Initiate first weight

The implementation of the training process begins with the process of giving the initial weight value. As explained in the design, the initial value of the weight is 1 divided by the number of features available. The code script can be seen in Figure 3.22.

```

public function _initializingWeight( $first = true ){
    if( $first ){
        for ($i=0; $i < (sizeof($this->featureInDB) +
        sizeof($this->featureOutDB)); $i++) {
            $this->weight[$i] = 1/(sizeof($this->featureInDB) +
            sizeof($this->featureInDB)); } return; }
        // get weight from db
        $weight = $this->weight_model->get();
        for ($i=0; $i < (sizeof($this->featureInDB) +
        sizeof($this->featureOutDB)); $i++) {
            $this->weight[$i] = $weight[$i]['weight']; }
        return;}

```

Figure 3. 23 Initiate First Weieht Script Code

2. Update weight function

Update weight is a function used to change the weight value if the result does not match the target. The modifying formula for weight values is explained in the research design section. The script code from the update weight can be seen in Figure 3.23.

```

// change value of weight
// Wbaru = wlama + target*x
for ($i=0; $i <= (sizeof($this->featureInDB) +
sizeof($this->featureOutDB) ); $i++) {
    if(isset($statistics[$i])){
        $this->weight[$i] = $this->weight[$i] +
        ($dataTarget[$index] * sizeof($statistics[$i])); } }

```

Figure 3. 24 Updating Weight Script Code

Example of calculation training process :

In process training, using data from questioners form. Data training can be seen in table 3.4.

Table 3. 4 Data.

No	Sentences	Target	Type
1	Alamat ayam asix malang dimn ya.....mau kesana ne	-1	cyberbully text
2	Muka edit 😊	1	non cyberbully text
3	Wjh ny brubah rubah ..😁tp stdkny dgn hijab lbh beuatiful😁	-1	cyberbully text
4	@ade_govinda numpang promo . cek ig kami 🤝👉 bagi kamu yg bosan kurus.. pengen banget gemuk ,😁😁😁 wajib coba suplemen ini 📄📄 dijamin manjur pasti naik 3-7 kg wa : 085733244789	-1	cyberbully text
5	wajanya kayak tante tante	1	non cyberbully text
6	mukanya boros	1	non cyberbully text
7	Muka nya boros bgt sii	1	non cyberbully text
8	gawenen ngopel coeg wkwkw	1	non cyberbully text
9	@lea_pezizir kalo sdh dijalani seluruh hukumannya baru bisa bebas, bgitu hukumnya, bego banget sih nih org #butahukum 🗝️🗝️	-1	cyberbully text
10	Bacod kntl	-1	cyberbully text
11	Terbang ke seluruh dunia dengan murah. Mau terbang dengan Pesawat Murah ? Trip menginap di hotel bintang harga murah ? Diskon up to 50berlaku seluruh Maskapai & Hotel, Domestik / Luar Negeri. Dijamin paling murah dari yg lain!! (Traveloka, PegiPegi, Agoda dan lainnya) No minimum payment / pax. DM/WA : 0821 9558 1130 #sheratonhotel #therinrahotel #shangriLahotel #swissbellinhotel #hotelclarion #marriotthotel #dbesthotel #senayanhotel #doubletreehotel #ibishotel #maskapai #lionair #garuda #sriwijaya #citylink #airasia #batikair #wingsair #flyemirates #qatarairways	-1	cyberbully text
12	Dan pada akhirnya fb qu nggak bisa nge hashtag di komenan .. Lanjutkan saudara2 qu 🤝🤝🤝 #TheVictoryOfPrabowo #QuickCountHoax #2019PrabowoSandi #2019GantiPresiden	-1	cyberbully text
13	@manado_bajualmurah saya ikutin... Anda coba ikutin lagi.. Sdkt saja saya	-1	cyberbully text

	jelaskan... ADP mentweet "saya, ludahi orang yg sdh membela penista Alquran "bgtlah twetnya kurang lebih dan idiot yg menghalangi utk berdemokrasi.. Dimn salahnya dan ADP menyebutkan namakah klu itu ujaran kebencian		
14	maklum anak kemarin soree yo	-1	cyberbully text
15	Blm mAndi yak om @duljaelani	-1	cyberbully text
16	Cuk!! Newbie!! Bleberan!!! Dasar anak muda!! Bwahahahahaha..	1	non cyberbully text
17	@vega.antares Maklum rekk... Newbie polll hahahahaha	-1	cyberbully text
18	Jangan cebok Dul, nanti pantatmu ungu	1	non cyberbully text
19	@vega.antares - AHAHAHAH!... amburadul si beDuL!..	-1	cyberbully text
20	@rrioricardo sya anak jaman old jg beleber td tintanya	-1	cyberbully text
21	Tumpeh tumpeh	1	non cyberbully text
22	Acak acakan bgt	-1	cyberbully text
23	Borosan orangnya sama tinta	-1	cyberbully text
24	@lea_pezizir bisa gitu yaa mbaa emangny	-1	cyberbully text
25	Aku nyoblos bapakmu cuk!!!\ud83e\udd18ben dadi anggota dpr ae, nek dadi musisi sakno model2 artis tik2 ilang endorse2 e....	-1	cyberbully text
26	Kapan si mamad ptong k..tl ni \ud83e\udd23\ud83e\udd23 serius Beta tanya ni \ud83d\ude4f\ud83d\ude4f	-1	cyberbully text
27	kamu dah makan njing	-1	cyberbully text
28	kamu udah makan anjing	-1	cyberbully text
29	asu banyak guguknya cuy	-1	cyberbully text
30	ah, masih bocah mirip babi	1	non cyberbully text
31	Banci kaleng ama duit mah cantik smua bisa di sulap hahaha	1	non cyberbully text
32	Beneran LBH suka MB Luna yg sekarang..... ♡♡	-1	cyberbully text
33	Kak lun dlu aku ga suka sama kak lun tapi semakin kesini kak luna semakin adem ga neko2 jd ngefens aku,, semakin cantik,, sukses terus ya kak...	-1	cyberbully text
34	Banci halu	1	non cyberbully text
35	Makin lama makin cantik luar dalam! Gak peduli masa lalu nya, y Pntng gak sombong dan baik	-1	cyberbully text

36	Kak lucinta ak ngefens sama kaka.. Apa sih rahasia cantiknya.. Pengen deh kaya kk	-1	cyberbully text
37	lumayan jg ni bencong	1	non cyberbully text
38	Tutup foto ini bagian hidung sampe mukutnyakliatan bgt wajah laki nya ...gk bs di boongin , jidat lo msh tllu lebar unt bs jd perempuan ..hahaha	-1	cyberbully text
39	Kk sbnarnya cwok atau cwek? Bnyk Brita ttg kk d TV n sosmed. Aq JD bngung. Aplgi AQ nton dyoutube ad ket nma kk operasi klamin. Smoga Brita itu gk bnr ya ka Lucinta	-1	cyberbully text
40	Ya allah si BENCONG nya nonggol lagi	1	non cyberbully text
41	@erih.riflan ganti otakmu bang	1	non cyberbully text
42	Tau gak klo ORANG BODOH ITU GAK BISA MIKIR???	1	non cyberbully text
43	Bibir nya itu lo mau tak cubit gemes	-1	cyberbully text
44	Bong coba deh mikir rasional pake otak sampingan itu cacian makian fitnah gajelas itu lu melek pake otak coba mikir dgn akal yg jernih coba kalo bisa ngaca ! Klo tetep gtu y mungkin emang udah cebong kuadrat kali ya	-1	cyberbully text
45	Malunya anak2mu mas, orang tua ko sukanya nebar kebencian, syukurlah mba maia cerai kalau gak pasti malu juga...	1	non cyberbully text
46	Horeeeee jokowi juara ,,,,,,hahahahahahaha ahaha mikiiiiir botak ah	-1	cyberbully text
47	@chahyaanirmala09 otakmu harus sering diupgrade biyar engga oleng mbae koeh cuh	-1	cyberbully text
48	Cebong tolol dan gak waras	1	non cyberbully text
49	Bisanya niru mulu tu cocotnya aja mrongos gitu	1	non cyberbully text
50	@ithaprimusklaou kebencian cuman milik cebong.. kita sih asik2 aja liat lu milih no 1, gak kaya lu benci klo liat org milih no 2..	-1	cyberbully text
51	Ya mmng tukang jiplak	-1	cyberbully text
52	Ehh tikus curuttttt, liat tu bibir nya gemeshhh	1	non cyberbully text
53	Klo gak ngibul ya plagiat, trus klaim milik orang,	-1	cyberbully text

54	ealah wong stress	-1	cyberbully text
55	Polower, trnyata otaknya pada kosong gak ada gebrakan gimana rakyat mau milih srul?	-1	cyberbully text
56	@virriyani_4 komen orang stres ya gini, ngawur ?	-1	cyberbully text
57	Muke gile	-1	cyberbully text
58	Goblok ee, disindir ga berasa	1	non cyberbully text
59	Wah ini koment tolol nih, itu lg disindir malah merasa hebat	-1	cyberbully text
60	@bieba_26 hahahaha sontoloyo.	-1	cyberbully text
61	E bong eh kecebong	-1	cyberbully text
62	rumah gue banyak ayam	-1	cyberbully text
63	komodo sama buaya beda ya	-1	cyberbully text
64	Cebong mana cebong... Hahaha	-1	cyberbully text
65	Dulu dihujat, dikatain, dicela... Pas berhasil ditiru	-1	cyberbully text
66	Justru postingan lu yg kyak gini yg bakal jatohin kubu loe sendiri q mau milih jdi males soale timsesnya kayak lo	-1	cyberbully text
67	Cebong no program, bisanya niru	-1	cyberbully text
68	@femmyjulia19 udah ga laku di musik jd jualan abab	-1	cyberbully text
69	Memang dasar CEBONG hidup di dua alam	-1	cyberbully text
70	@virriyani_4 ealah ealah cebong tolol	-1	cyberbully text
71	Sy yakin anda nglakuin ini karna ego anda sangat besar' hati anda sebenarnya lebih berat ke No 1 ketimbang No 2	-1	cyberbully text
72	Biarlah...jadi jariaah yaa	-1	cyberbully text
73	Nahhh gitu kan bagus, itu namanya orang yg jujur. jujur akui kalau program Anies dan Sandi itu keren, dan bmnfaat utk rakyat...	-1	cyberbully text
74	Demi last minute pencitraan apapun ditabrak tapi rakyat di arus bawah sudah pd melek ga bisa di bohongin lg	-1	cyberbully text
75	Kirain lagu doang yg dicover, ternyata program yg lagi hits juga bisa dicover	-1	cyberbully text
76	JENDERAL KARDUS, BADAN GEDE SEPERTI BAGONG, BELAGAT SONGONG OMONG KOSONG.	1	non cyberbully text
77	Apasih yg tdk di copy paste,	-1	cyberbully text
78	aku punya anjing	-1	cyberbully text
79	lo anjing	1	non cyberbully text

80	kambing kamu lucu	-1	cyberbully text
81	anjing	-1	cyberbully text
82	yah ada jablay nih, aslinya dia laki loh	1	non cyberbully text
83	dasar homo	1	non cyberbully text
84	cocotmu mantap	-1	cyberbully text
85	Bencong	1	non cyberbully text
86	Kmu bencong	1	non cyberbully text
87	Jakun nya kelihatan bngt. Kenapa ya? Apakah memang gk bisa dihilangkan. Meski dgn cara operasi ? Soalnya aku tu sering ngaceng kalo lihat dia.. Tapi off lagi stiap liat jakunnya itu.. 🤔	-1	cyberbully text
88	Kak @lucintaluna kok muka nya lama lama mirip cowok sih kak ??	-1	cyberbully text
89	Bisa untuj muka juga?	-1	cyberbully text
90	Mikir banci	-1	cyberbully text
91	Bencong bercangkang	1	non cyberbully text
92	VIRAL!!!!!!! 🤔🤔🤔🤔 RAHASIA PUTIH GIGI ARTIS & SELEBGRAM GiGI ANDA KUNING? PLAK? ADA KARANG GIGINYA? BIKIN KAMU GAK PERCAYA DIRI GOSOK GIGI MU DENGAN SUPER WHITENING TEETH READYSTOCK PAKET PEMUTIH & PERAWATAN GIGI SUPER WHITENING TEETH ASLI 100RIGINAL HANYA TINGGAL 4 PAKET TERAKHIR . BERIKUT FUNGSI SUPER WHITENING TEETH ORIGINAL : 1. Memutihkan gigi dengan cepat 2. Menghilangkan noda & karang gigi 3. Menyegarkan nafas & mencegah bau mulut 4. Menguatkan gigi mencegah gigi berlubang 5. Menjaga kesehatan gigi & mencegah sekaligus mengobati gusi berdarah 6. Cocok untuk pengguna behel karena terbuat dari bahan alami tanpa bahan kimia yang membuat korosi behel untuk informasi seputar produk silahkan Follow: @pemutih_gigi79 ORDER / KONSULTASI : WA : 0821-1993-1594 FAST(RESPON)	-1	cyberbully text
93	Banci banci banci banci	1	non cyberbully text
94	Assalamualaikum ga jawab banci :v	-1	cyberbully text

95	Awalnya aku emang nyangka lucinta itu perempuan. Tapi pas diperhatiin mukanya kayak banci, banyak jd yg viral viral sekarang tentang lucinta itu laki laki. Dan sekarang fix, menurut gue lucinta itu laki laki	-1	cyberbully text
96	Kntl	1	non cyberbully text
97	Bencong kesasar apa yah?	1	non cyberbully text
98	Ni bencong ya	-1	cyberbully text
99	Opsi 15	-1	cyberbully text
100	Oi bencong	1	non cyberbully text
101	Dulu lu suka di tusuk pantat lu kan	-1	cyberbully text
102	Pasti keluar tai ya	-1	cyberbully text
103	Dasar bencong	1	non cyberbully text
104	Bencong Anjing	1	non cyberbully text
105	Opsi 21	-1	cyberbully text
106	Banci kontolllllllllll	1	non cyberbully text
107	Bencong / wanita berjagun wkwkwkwkwkw Uda kehabisan malu iyaaa .pansos mullu Bencong 🤪 @lucintaluna	-1	cyberbully text
108	@abdullahfajar.r lo malah bikin dosa sendiri tolol wkwk 🤪	-1	cyberbully text
109	Jual jam tangan panerai authentic, tas braun buffel authentic....	-1	cyberbully text
110	@ainnisalsabila bocah bacot bgst	-1	cyberbully text
111	mana yang katanya rambutnya ditiru sama blackpink? NGAREP LU BGST	-1	cyberbully text
112	@ainnisalsabila bocah curut jembut aja blum numbuh udh bnyk bacot aja!!! ingus tu di lap anjng 🤪	-1	cyberbully text
113	Mas fatah ngewe yuuu 🤪	-1	cyberbully text
114	Selamat siang buat para pecinta gadget, VIVO selalu menghadirkan sesuatu yang terbaru sesuai impian dan keinginan anda semua dengan harga yang sesuai poket kalian guys 🤪 perkenalkan hanya di GALTEKINDO semua impian anda bisa jadi kenyataan seperti beli cash hp, CREDIT PAKAI DP juga Bisa CREDIT TANPA Dp, silahkan mohon budayakan membaca lokasi kita di TANGERANG CIPONDOH Syaratnya apa aja sih?? Beghhh gampang banget pokoknyaaa!! 1. Jika bossnya kerja = KTP + SLIP GAJI, 2. Jika bossnya usaha = KTP +	-1	cyberbully text

	<p>SKU. GAMPANG BANGET KANNN TUH!!! KREDIT HP DI KAMI MAHH GAMPANG BANGETTT NGETTT NGETTTT.. PENGAJUANNYA bisa via WHATSAPP doang!! Ga perlu cape2, jauh2, buang2 bensin kemariii! TAPI jika sudah diACC, ambil hpnya kudu kemari ya boss2ku, biar pake HP BARUnya lebih afdollll gitu.. 😊</p> <p>BENER2 TELEK DIBUKA DIDEPAN MATA BOSSNYA.. 😊😊😊 note : khusus wilayah JABODETABEK yaa boss.. Bagi yg mau ajuin : bisa langsung WA ke 085977304969 Laela Terima kasih, GALTEKINDO.</p>		
115	Itu..blg puas ...kyk nging kuping gua	-1	cyberbully text
116	Lo banci ya....	1	non cyberbully text
117	Sabunnya aman dipake coli kak???	-1	cyberbully text
118	<p>Dijual Toyota Agya TRD 1.2 A/T 2018 Semua spek original, tidak ada yang diganti Warna silver, NO BARET KM 10rb, plat exp 2023, pajak exp 26/09/19 Plat B DKI Jakarta Ban masih bagus semua ga pernah meletus ataupun bocor, sensor parkir nyala, wiper berfungsi normal, lampu all normal. Interior bersih (pakaian wanita), AC dingin, radio bluetooth + usb bisa buat telponan juga, audio mantap, matic responsif, kaki2 bagus. Pernah dipakai di dalam kota saja, tidak ke luar kota, garansi llumar 2x, track record servis a2000 lengkap, bpkb lengkap, stnk lengkap, pajak lengkap, faktur juga lengkap. Free pelayanan emergency a2000 5 tahun HARGA OTR 135,000,000 NO MAKELAR, SAYA GA SUKA. Harga segitu biasanya masih ada overkredit atau harga khusus kredit dealer. SAYA LUNAS, GA ADA UTANG, JANGAN TEGA. Trims. WA 081380742456</p>	-1	cyberbully text
119	Makin hari makin cantik aja kak	-1	cyberbully text
120	Sini sama abang wkwk	-1	cyberbully text
121	bangga si bencong di puja puja didunia,dia bakal nyesel akhir hidupnya bagaimana	1	non cyberbully text
122	Salfok amang jidat si om fatah	-1	cyberbully text

123	Ganteng Lucinta	-1	cyberbully text
124	hai tante ganteng banget sih	-1	cyberbully text
125	Ihhh cantik kayak babi	1	non cyberbully text
126	Kok pake baju cewe bukannya pake baju cowo entar kalo pake baju cewe di kata benchong lagi	-1	cyberbully text
127	aura maco lo gakan pernah ilang samsudin.	1	non cyberbully text
128	@mingue.l@pyashinta tembak cyn tembaaakkkk	-1	cyberbully text
129	Subhalallah cantiknya luar biasa	-1	cyberbully text
130	banci	1	non cyberbully text
131	Tambah cantik mb lun	-1	cyberbully text
132	Aku suka lipstiknya. Kak pake lipstik apa itu. Aku suka warnanya	-1	cyberbully text
133	Nda cantik le kong lonte liar tabiar	1	non cyberbully text
134	@lucintaluna yes yes yes #1 supporter kak! much love ♡♡♡♡♡ sukses selalu cc #EnjoyYourLife	-1	cyberbully text
135	Ka luna gendutan ya @lucintaluna	-1	cyberbully text
136	Kk mkin cntikk suka liatny	-1	cyberbully text
137	Laki amat mukanya	-1	cyberbully text
138	@ainnisalsabila banyak bacod lu kaga usah ikut ikutan bocah ingusan	-1	cyberbully text
139	@ainnisalsabila dasar babi laknat	1	non cyberbully text
140	@ainnisalsabila bocil sok nasehatin..... Ujung2nya minta dikepoin.. Dasar kon.tol	1	non cyberbully text
141	Kontolnya kak, silahkan diorder... Siapa tau mau balik lagi kek dulu, kalo minat DM ya kak nanti aku kasih kontol kuda untuk mu @lucintaluna 🐎	1	non cyberbully text
142	@ainnisalsabila BACOD	-1	cyberbully text
143	KALAU YANG PROMOSI WARIA, BENCONG, TRANSGENDER KAYAK GINI... JADI RAGU MAU BELI PRODUKNYA	-1	cyberbully text
144	@besty_ayu norak,, duh maaf pengikutnya ratu salome cong marah ... 🤔	-1	cyberbully text
145	Hai Cowok	-1	cyberbully text
146	Cowok	-1	cyberbully text
147	Kq kayak laki ya ???	-1	cyberbully text
148	Wanita berktl :v	1	non cyberbully text
149	Salah satu waria tercantik di Indonesia	-1	cyberbully text
150	NGOMONGNYA MENDESAH @lucintaluna wkakakkakaka	-1	cyberbully text

151	Hai Ganteng	-1	cyberbully text
152	Bisul apa jakun ?	-1	cyberbully text
153	Ada yg nyangkut di tenggorokan	-1	cyberbully text

Training process is elaborated as follow:

1. Get feature of data training. some result can be explained in table 3.5.

Table 3. 5 Example Feature Extraction

No	Sentence	Statistics
0	Banci kaleng ama duit mah cantik smua bisa di sulap hahaha	{"1":["banci"],"7":0.090909090909091,"totalBully":1}
1	Beneran LBH suka MB Luna yg sekarang..... ????	{}
2	Kak lun dlu aku ga suka sama kak lun tapi semakin kesini kak luna semakin adem ga neko2 jd ngefens aku,, semakin cantik,, sukses terus ya kak...	{}
3	Banci halu	{"1":["banci"],"7":0.5,"totalBully":1}

2. First weight value is $1/\text{feature}$, so the weight is $1/7$.and first margin is 0.
3. Calculate value of first sentences is:

$$a. \quad x_1 = 0, x_2 = 1, x_3 = 0, x_4 = 0, x_5 = 0, x_6 = 0, x_7 = 0,$$

$$x_8 = 0.090909090909091$$

- b. Result calculation using hyperplane function 3.6 is :

$$\begin{aligned} \text{value} &= \frac{1}{7} * 0 + \frac{1}{7} * 1 + \frac{1}{7} * 0 + \frac{1}{7} * 0 + \frac{1}{7} * 0 + \frac{1}{7} * 0 + \frac{1}{7} * 0 \\ &+ \frac{1}{7} * 0.090909090909091 \end{aligned}$$

Result is 0.15584416.

- c. If result < 0 , the text as non cyberbullying, and the opposite as cyberbullying text. Because 0.15584416 is more than 0. So type and target is suitable. And process is calculate of second sentences.
- d. If type result and target is not suitable, weight is updated. The example is:

$$w_{2 \text{ new}} = \frac{1}{7} + -1 * 0.15584416$$

$$w_{8 \text{ new}} = \frac{1}{7} + -1 * 0.15584416$$

- e. Check all how much correct data and not correct as temp acuration.
 - f. Then loop text util end process a-c. and check the now accuration.
 - g. If now accuration $<$ low accuration, process is next text. And if accuration $>$ low accuration, weight is updated.
4. Process is looping until end of text with the best separation data.

For the complete process of training, it can be seen from the implementation of the training process with the script code contained in the Figure 3.24 and Figure 3.25.

```

public function training ( $request, $response, $args ) {
    $this->_index();
    $this->params = $request->getParsedBody();
    $this->args = $args;
    // get number data training from db from user request
    $intTraining = $this->params['training'];
    $intTarget = $this->params['target'];
    $this->initModel("research");
    // get data training and data target from db
    $data = $this->research_model->get();
    $data = $this->_loopWord($data);
    // initial data training
    $dataTraining = $data[$intTraining]['text'];
    // initial data target of training
    $dataTarget = $data[$intTarget]['text'];
    // multiple preprocessing
    $data = $this->_loopPreprocessing( $dataTraining );
    // loop search in text
    $data = $this->_loopSearchWord( $dataTraining, $data );
    // set first weight
    $this->_initializingWeight(true);
    // loop training
    $this->_loopTraining( $data, $dataTarget );
    // save weight to db
    $this->_loopSaveWeight( );
    // testing of training
    $data = $this->_loopTesting( $data );
    // get accuration

```

Figure 3. 25 Training Script Code Part.1

```

$this->accuration = $this->_checkAccuration( $data,
$dataTarget );

// generate hyperplane
$hyperplane = "";
for ($i=0; $i < (sizeof($this->weight)); $i++) {
    if( $i < sizeof($this->weight) - 1){
        $hyperplane .= $this->weight[$i] ." * ".
"x".($i+1)." + ";
    }else{
        $hyperplane .= $this->weight[$i] ." * ".
"x".($i+1);}
}
// Return a Feedback
$this->db = null; return $response->withJSON( array("status" => true,
"accuration" => ($this->accuration*100),"hyperplane" => $hyperplane,
"data" => $data) );}

```

Figure 3. 26 Training Script Code Part.2

3.2.13 Testing Implementation

Implementation of the testing process by writing a script to loop testing data to test using the hyperplane function that has been obtained in the training process. Here is an example of a manual testing process in the program.

Example text for testing is : “ihh cantik kayak babi”

Example full process calculation process at text to detect text type is: .

The result get from training process is:

$$\text{margin} = 0.038961038961039$$

$$\begin{aligned}
 \text{Hyperplane} &= 0.071428571428571 * x_1 + 1.0714285714286 * x_2 \\
 &+ 0.071428571428571 * x_3 + 0.071428571428571 * x_4 \\
 &+ 0.071428571428571 * x_5 + 0.071428571428571 * x_6 \\
 &+ -1.9285714285714 * x_7 + 0.071428571428571 * x_8
 \end{aligned}$$

So with the hyperplane, calculation can be explained as follow :

1. Value of feature is:

$$x_1 = 1, x_2 = 0, x_3 = 0, x_4 = 0, x_5 = 0, x_6 = 0, x_7 = 0, x_8 = 0,25$$

2. so, calculation by formula above is:

$$\begin{aligned} \text{value} &= 0.071428571428571 * 1 + 1.0714285714286 * 0 \\ &+ 0.071428571428571 * 0 + 0.071428571428571 \\ &* 0 + 0.071428571428571 * 0 \\ &+ 0.071428571428571 * 0 + -1.9285714285714 \\ &* 0 + 0.071428571428571 * 0.25 \end{aligned}$$

3. result of calculation above is 0.08928571
4. if result < margin, type text is cyberbullying, and if opposite the type is noncyberbullying. Because 0.08928571 is more than 0.038961038961039, the type is cyberbullying text.

Implementation of testing has done by write a script code. Script code of testing program is illustrated in Figure 3.26 and Figure 3.27.

```

public function knownTesting( $request, $response, $args ){
    $this->_index();
    $this->params = $request->getParsedBody();
    $this->args = $args;
    // get number data training from db from user request
    $intData = $this->params['id_data'];
    //set start time
    $rustart = gettrusage();
    $dataDB = $this->_insertData( $intData );
    // decode json
    $dataPre = json_decode($dataDB['text']);
    // multiple preprocessing
    $data = $this->_loopPreprocessing( $dataPre );
    // loop search in text
    $data = $this->_loopSearchWord( $dataPre, $data );
    // set first weight
    $this->_initializingWeight(false);
    // testing full
    $data = $this->_loopTesting( $data );
    // end Time
    $ru = gettrusage();
    $dataBully = $this->_getTextBully( $data );
    $dataResult = $this->_listResult( $data );
    // measure accuration (precision, recall)
    $dataTarget = json_decode($dataDB['target']);
    $value = $this->_measureAccuration( $dataTarget,
    $dataResult );
    // get data result from db
    $saved = $this->result_model->get( $intData );
    $action = "";
    $data = array(
        "data_id" => $intData,
        "result" => json_encode($dataResult),
        "precision" => $value['precision'],
        "recall" => $value['recall'],
        "speed" => $this->_runtime($ru, $rustart, "utime")
    );

    // check if has saved in db or no if yes update if no insert
    if( $saved == false ){
        // found in db
        // insert in db
        $action = $this->result_model->insert( $data );
    }else{
        // update in db
        $action = $this->result_model->update( $data );}
    $data['cyber_total'] = $dataDB['cyber_total'];
    $data['noncyber_total'] = $dataDB['noncyber_total'];
    $data['size'] = sizeof($dataPre);
    $data['text'] = $dataDB['text'];
    $data['target'] = $dataDB['target'];
    $this->db = null; return $response->withJSON( array("status"
=> $action, "data" => $data) );}

```

Figure 3.49 Testing Script Code Part.1

```

public function _loopTesting( $data, $index = 0){
    $resultData = 0;
    $statistics = $data[$index]['statistics'];

    for ($i=0; $i < (sizeof($this->featureInDB) + sizeof($this->featureOutDB)); $i++) {
        if(isset($statistics[$i])){
            $resultData += $this->weight[$i] *
                sizeof($statistics[$i]);}
        $result = ($resultData <= 0) ? -1 : 1;
        if( $result == -1 ){
            $data[$index]['result']['type'] = 'Positive Sentence';
            $data[$index]['result']['value'] = $resultData;
            $data[$index]['result']['text'] = $data[$index]['text'];
        }else{
            $data[$index]['result']['type'] = 'CyberBullying
Sentence';
            $data[$index]['result']['value'] = $resultData;
            $data[$index]['result']['text'] = $this->replaceWord(
                $data[$index]['text'], $data[$index]['statistics'] );}

        if( $index == sizeof($data) - 1 ) return $data;
        return $this->_loopTesting( $data, $index + 1 );}

```

Figure 3. 27 Testing Script Code Part.2

3.2.14 Design Interface Impelementation

The application that was built was named `bully_checker`. This application is built web-based using the PHP programming language and Code Igniter framework. Making the system interface uses HTML5, JavaScript and JQuery languages. Then to enhance the appearance of the web interface, the CSS Bootstrap framework is used. This application consists of four main pages, namely Home, Auth (Login and Register), Training, Testing, Docs, and Dictionary.

1. Home

Display of the home page as the default display when first entering the program. This page has an input that is used to fill in the text that will be detected. The home page before the detection process can be seen in Figure 3.27.

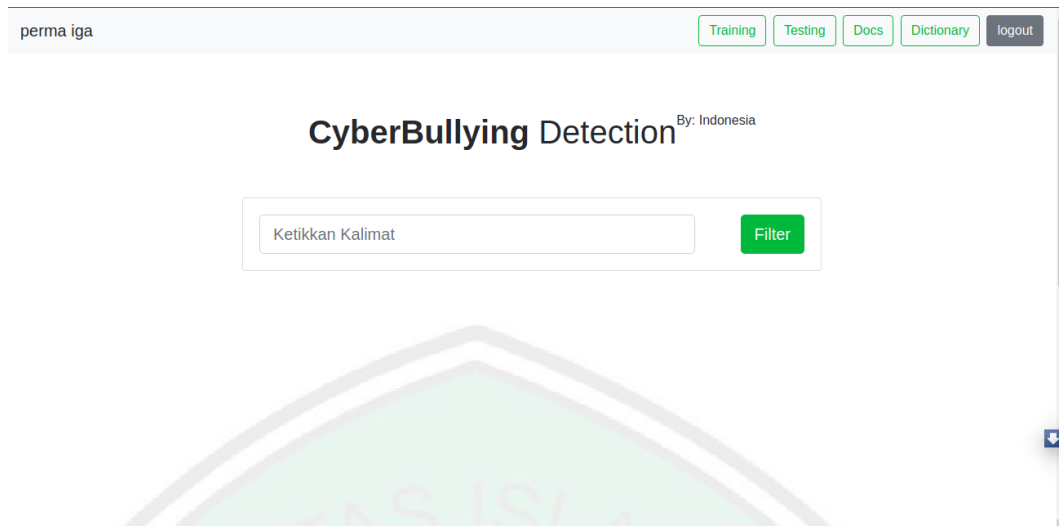


Figure 3. 28 Display of Home Page

After a text is inputted, the system will contact the API that I have created, which is to detect single texts. And the appearance of the time the user has got the results can be seen in Figure 3.28.



Figure 3. 29 Display of Home Page with Result

2. Auth

- Login

On the login page, there is an email input and password that was registered previously. The authentication function is to be able to

access the API. Display of login page can be seen in Figure 3.29.

Figure 3. 30 Display of Login Page

- Register

On the register page there is a full name input to fill in the full name, then email, and password. There is a register button for the registration process. Register page can be seen in Figure 3.30.

Figure 3. 31 Display of Register Page

3. Training

Training pages can only be accessed by registered users. In the case of safe training there is a data option, used to select data to be trained. Then there is the training button to do the next process. Training Display can be seen in Figure 3.31.

perma iga Training Testing Docs Dictionary Logout

Pilih Data:
training c=83 nc=75

training

Result :

margin = 0.071428571428571

hyperplane = -0.92857142857143 * x1 + 1.0714285714286 * x2 + 1.0714285714286 * x3 + 0.071428571428571 * x4 + -0.92857142857143 * x5 + 1.0714285714286 * x6 + -0.92857142857143 * x7 + -0.92857142857143 * x8

No	Sentence	Statistics	Result	
			Type	Value
0	bgtulah kalo tikus terpojok.... nabrak sana	{"0":["tikus"],"7":0.14285714285714,"totalBully":1}	Positive Sentence	-1.0612244897959

Figure 3. 32 Display of Training Page

If you have selected data and processed it, the data will come out as shown in the picture.

4. Testing

Testing pages can only be accessed by registered users. In the case of safe training there is a data option, used to select the data to be tested. Then there is the training button to do the next process. Testing page can be seen in Figure 3.32.

perma iga Training Testing Docs Dictionary logout

Pilih Data:

testing c=15 nc=35

testing

Result :

total bully = 15
total non bully = 35
precision = 0.40740740740741
recall = 0.7333333333333333
speed = 48

Figure 3. 33 Display of Testing Page

If you have selected data and processed it, the data will come out as shown in the picture.

5. Docs

On the docs page, this page aims to display steps to be able to access the API as well as the url that can be accessed. The url contains tokens from each user to be able to use as needed. Docs page can be seen in Figure 3.33.

Welcome to CyberBullying Detection API

Berikut Ini Adalah API url untuk menggunakan API Cyberbullying detection

API URL METHOD HTTP : POST

localhost/cyber-detection/v1/f11tras1/a097e468111fb1edfe2e936f1ad5d249

Params :

text : {text yang ingin di dideteksi}

Page rendered in 0.0576 seconds Cyberbullying Detection v 1.0

Figure 3. 34 Display of Docs Page

6. Dictionary

This page displays, what word do I use the system to detect texts. The purpose is to display this page so that the user can see the list of words in this system. Every type of word cyberbully is shown here. Dictionary page can be seen in Figure 3.34.

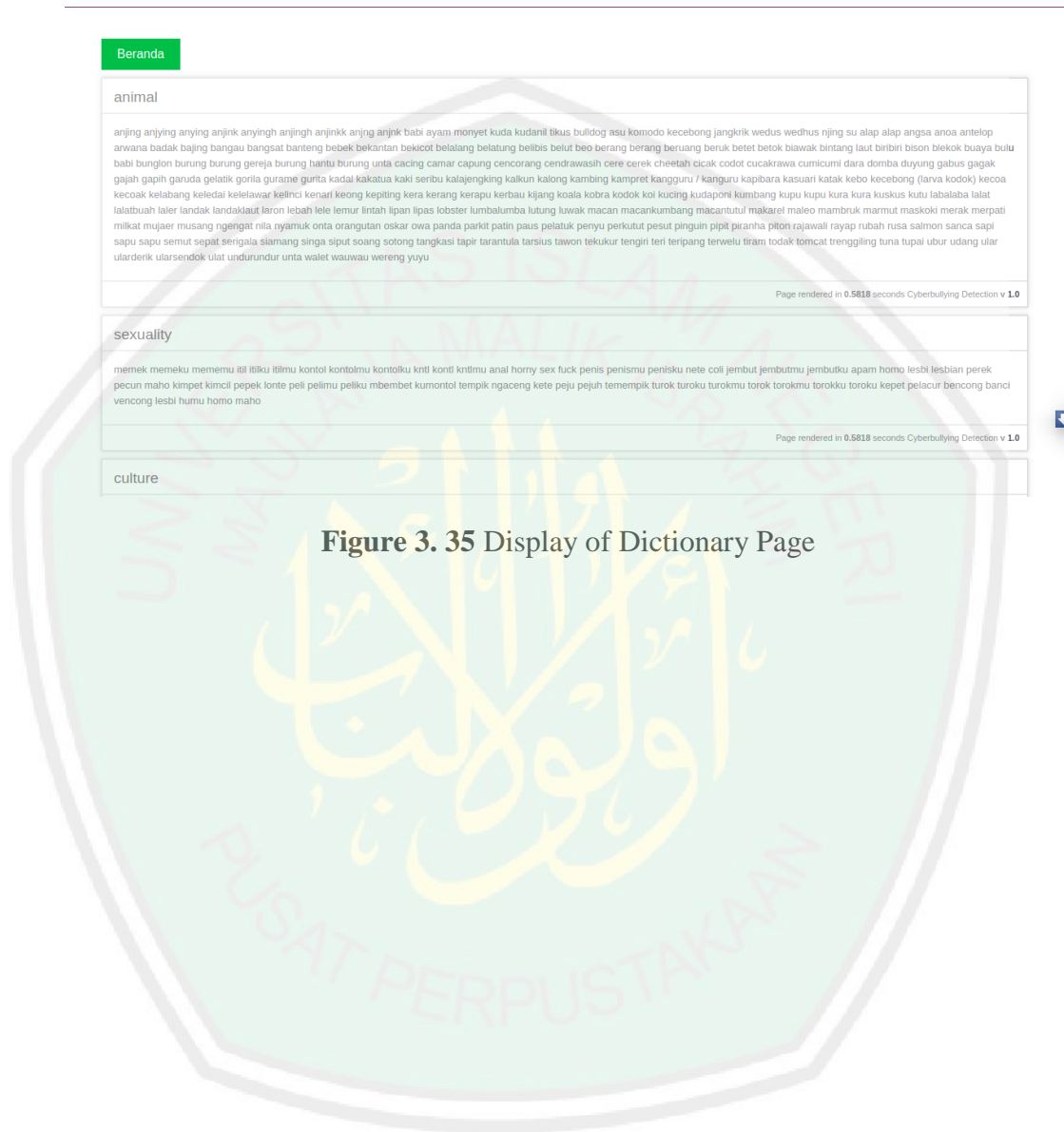


Figure 3. 35 Display of Dictionary Page

CHAPTER IV

RESULT AND ANALYSIS

4.1 Experimental Setup

Experimental setup is a step to manage how the experiment will be conducted. The number of trials is adjusted to the amount of research question. There are three research questions, namely measure accuracy, measure speed, and reliable measure. In this sub-chapter also explained the data used in testing. The descriptions are as follows:

1. Measuring SVM accuracy for detecting cyberbully textuuals.

According to Sokolova (2009), in binary classification, measurement of accuracy can be done by calculating the precision and recall, and referring to the confusion matrix table. The confusion matrix table is as follows table 4.1:

Table 4. 1 Confussion Matrix

Class	Positive Classificated	Negative Classificated
Positif	TP (True Positive)	FN (False Negative)
Negatif	FP (False Positive)	TN (True Negative)

Meanwhile, the applicable formula is as follows:

$$Precision = TP / (TP + FP) \quad (4.1)$$

$$Recall = TP / (TP + FN) \quad (4.2)$$

Precision is the level of accuracy between the information requested by the user and the answer given by the system. Whereas recall is the success rate of the system in rediscovering information. to measure the accuracy of SVM, valid data used which obtained by survey in the previous step. Our data is for 2 classes is data

labeled cyberbullying and non cyberbullying. Then the data is tested all, and calculated the precision value and recall value.

2. Measuring SVM speed to detect textual cyberbullying

According to Wirahutama (2011), the speed of a software can be measured by finding the average value of speed during program execution. So that in this study, to measure the speed of the SVM algorithm, the program executes has 3 scenario with total data different each scenario is 40,72, and 151 data, and try it 10 times and calculates the average time to process data and return result. The formula is as follows:

$$T_{avg} = \sum_{i=1}^{10} ti / 10 \quad (4.3)$$

Where T_{avg} is average time and ti is time each process.

Then to get speed, is using time divide by total data.

$$V = \frac{n}{T_{avg}} \quad (4.4)$$

Where, v is speed, T_{avg} is time average, and n is number of data in process.

3. Measuring system reliability

According to Pan (1999), software reliability is the probability of operating errors over a certain period in a specific environment with the following formula:

$$p = \frac{\text{case failure}}{\text{total time}} \quad (4.5)$$

Where, p is probability, case failure is total error, and total time is time at testing.

According to Wirahutama (2011), Reliability can be seen from the system being able to serve the needs of users without problems that can disrupt the user's

comfort in using information systems. If refer to the above explanation, in measuring the system reliability, demo program is spread to try by the user. Then calculate the probability by using input data entered by the user by calculating the correct and incorrect output according to the user. As an example of calculation, if there are 100 people who try and 25 person get error, then the probability error is $25/100$ which is 2.5% probability error.

The data for testing is collected from the questioner process described in chapter 3. The data can be seen in table 4.2:

Table 4. 2 Data Testing

Number	Sentence Number	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
1	1 to 31	Testing	Training	Training	Training	Training
2	32 to 62	Training	Testing	Training	Training	Training
3	63 to 93	Training	Training	Testing	Training	Training
4	94 to 124	Training	Training	Training	Testing	Training
5	125 to 153	Training	Training	Training	Training	Testing

4.2 Experimental Result

In this sub-section, the results obtained from the trial process are described in section 4.1. but the results consist of several tables, with each research question point having a table and the results of each. The explanation will be explained based on the previous test design. The results of the test are as follows:

1. Accuracy measurement

Accuracy measurement, carried out on data that has become the output of the system. The data in question is training data found in sub-section 4.1. the data is processed on the system, then outputs. The output can be seen in table 4.3.

Table 4. 3 Testing Result

Sentences Number	Sentences	Value
1	Alamat ayam asix malang dimn ya.....mau kesana ne	-0.1696
2	Muka edit ?	-1.5714
3	Wjh ny brubah rubah ..?tp stdkny dgn hijab lbh beuatiful?	-0.1214
4	@ade_govinda numpang promo . cek ig kami ?? bagi kamu yg bosan kurus.. pengen banget gemuk ,??? wajib coba suplemen ini ?? dijamin manjur pasti naik 3-7 kg wa : 085733244789	-0.9908
5	wajanya kayak tante tante	0
6	mukanya boros	0
7	Muka nya boros bgt sii	-1.3143
8	gawenen ngopel coeg wkwkw	-1.4107
9	@lea_pezizir kalo sdh dijalani seluruh hukumannya baru bisa bebas, bgiu hukumnya, bego banget sih nih org #butahukum ??	-1.0357
10	Bacod kntl	-2.7857
11	Terbang ke seluruh dunia dengan murah. Mau terbang dengan Pesawat Murah ? Trip menginap di hotel bintang harga murah ? Diskon up to 50berlaku seluruh Maskapai & Hotel, Domestik / Luar Negri. Dijamin paling murah dari yg lain!! (Traveloka, PegiPegi, Agoda dan lainnya) No minimum payment / pax. DM/WA : 0821 9558 1130 #sheratonhotel #therinrahotel #shangriLahotel #swissbellinhotel #hotelclarion #marriotthotel #dbesthotel #senayanhotel #doubletreehotel #ibishotel #maskapai #lionair #garuda #sriwijaya #citylink #airasia #batikair #wingsair #flyemirates #qatarairways	0.04501
12	Dan pada akhirnya fb qu nggak bisa nge hashtag di komenan .. Lanjutkan saudara2 qu ??? #TheVictoryOfPrabowo #QuickCountHoax #2019PrabowoSandi #2019GantiPresiden	-1.025
13	@manado_bajualmurah saya ikutin... Anda coba ikutin lagi.. Sdkt saja saya jelaskan... ADP mentweet "saya, ludahi orang yg sdh membela penista Alquran "bgtlah twetnya kurang lebih dan idiot yg menghalangi utk berdemokrasi.. Dimn salahnya dan ADP menyebutkan namakah klu itu ujaran kebencian	0.02439
14	maklum anak kemarin soree yo	0
15	Blm mAndi yak om @duljaelani	0
16	Cuk!! Newbie!! Bleberan!!! Dasar anak muda!! Bwahahahahahaha..	0
17	@vega.antares Maklum rekk... Newbie polll hahahahaha	0
18	Jangan cebok Dul, nanti pantatmu ungu	0

19	@vega.antares - AHAHAHAH!... amburadul si beDuL!..	0
20	@rioricardo sya anak jaman old jg beleber td tintanya	0
21	Tumpeh tumpeh	0
22	Acak acakan bgt	0
23	Borosan orangnya sama tinta	0
24	@lea_pezizir bisa gitu yaa mbaa emangny	0
25	Aku nyoblos bapakmu cuk!!!\ud83e\udd18ben dadi anggota dpr ae, nek dadi musisi sakno model2 artis tik2 ilang endorse2 e....	0
26	Kapan si mamad ptong k..tl ni \ud83e\udd23\ud83e\udd23 serius Beta tanya ni \ud83d\ude4f\ud83d\ude4f	0
27	kamu dah makan njing	-0.4107
28	kamu udah makan anjing	-0.4107
29	asu banyak guguknya cuy	-0.4107
30	ah, masih bocah mirip babi	-0.3143
31	Banci kaleng ama duit mah cantik smua bisa di sulap hahaha	-0.1039
32	Beneran LBH suka MB Luna yg sekarang..... ????	0
33	Kak lun dlu aku ga suka sama kak lun tapi semakin kesini kak luna semakin adem ga neko2 jd ngefens aku,, semakin cantik,, sukses terus ya kak...	0
34	Banci halu	0.10714
35	Makin lama makin cantik luar dalam! Gak peduli masa lalu nya, y Pntng gak sombong dan baik	0
36	Kak lucinta ak ngefens sama kaka.. Apa sih rahasia cantiknya.. Pngen deh kaya kk	0
37	lumayan jg ni bencong	0.58929
38	Tutup foto ini bagian hidung sampe mukutnyakliatan bgt wajah laki nya ...gk bs di boongin , jidat lo msh tllu lebar unt bs jd perempuan ..hahaha	-2
39	Kk sbnarnya cwok atau cwek? Bnyk Brita ttg kk d TV n sosmed. Aq JD bngung. Aplgi AQ nton dyoutube ad ket nma kk operasi klamin. Smoga Brita itu gk bnr ya ka Lucinta	0
40	Ya allah si BENCONG nya nonggol lagi	0.79592
41	@erih.riflan ganti otakmu bang	0
42	Tau gak klo ORANG BODOH ITU GAK BISA MIKIR???	-0.1429
43	Bibir nya itu lo mau tak cubit gemes	-2.3393
44	Bong coba deh mikir rasional pake otak sampingan itu cacian makian fitnah gajelas itu lu melek pake otak coba mikir dgn akal yg jernih coba kalo bisa ngaca ! Klo tetep gtu y mungkin emang udah cebong kuadrat kali ya	0

45	Malunya anak2mu mas, orang tua ko sukanya nebar kebencian, syukurlah mba maia cerai kalau gak pasti malu juga...	0
46	Horeeeee jokowi juara ,,,,,,hahahahahahaha ahaha mikiiiiir botak ah	0
47	@chahyaanirmala09 otakmu harus sering diupgrade biyar engga oleng mbae koeh cuh	0
48	Cebong tolol dan gak waras	-1.6286
49	Bisanya niru mulu tu cocotnya aja mrongos gitu	0
50	@ithaprimusklaui kebencian cuman milik cebong.. kita sih asik2 aja liat lu milih no 1, gak kaya lu benci klo liat org milih no 2..	0
51	Ya mmng tukang jiplak	0
52	Ehh tikus curuttttt, liat tu bibir nya gemeshhh	-0.3393
53	Klo gak ngibul ya plagiat, trus klaim milik orang,	0
54	ealah wong stress	0
55	Polower, trnyata otaknya pada kosong gak ada gebrakan gimana rakyat mau milih srul?	0
56	@virriyani_4 komen orang stres ya gini, ngawur ?	0
57	Muke gile	0
58	Goblok ee, disindir ga berasa	-0.3143
59	Wah ini koment tolol nih, itu lg disindir malah merasa hebat	-1.1039
60	@bieba_26 hahahha sontoloyo.	0
61	E bong eh kecebong	0.58929
62	rumah gue banyak ayam	0.58929
63	komodo sama buaya beda ya	-0.2286
64	Cebong mana cebong... Hahaha	0
65	Dulu dihujat, dikatain, dicela... Pas berhasil ditiru	0
66	Justru postingan lu yg kyak gini yg bakal jatohin kubu loe sendiri q mau milih jdi males soale timsesnya kayak lo	0
67	Cebong no program, bisanya niru	0
68	@femmyjulia19 udah ga laku di musik jd jualan abab	0
69	Memang dasar CEBONG hidup di dua alam	0
70	@virriyani_4 ealah ealah cebong tolol	-1.1143
71	Sy yakin anda nglakuin ini karna ego anda sangat besar' hati anda sebenarnya lebih berat ke No 1 ketimbang No 2	0
72	Biarlah...jadi jariaah yaa	0
73	Nahhh gitu kan bagus, itu namanya orang yg jujur. jujur akui kalau program Anies dan Sandi itu keren, dan bmnfaat utk rakyat...	0
74	Demi last minute pencitraan apapun ditabrak tapi rakyat di arus bawah sudah pd melek ga bisa di bohongin lg	0
75	Kirain lagu doang yg dicover, ternyata program yg lagi hits juga bisa dicover	0

76	JENDERAL KARDUS, BADAN GEDE SEPERTI BAGONG, BELAGAT SONGONG OMONG KOSONG.	0
77	Apasih yg tdk di copy paste,	0
78	aku punya anjing	-0.2381
79	lo anjing	-0.3929
80	kambing kamu lucu	-1.4762
81	anjing	-0.8571
82	yah ada jablay nih, aslinya dia laki loh	-1.0446
83	dasar homo	-0.3929
84	cocotmu mantap	0
85	Bencong	-0.8571
86	Kmu bencong	-0.3929
87	Jakun nya kelihatan bngt. Kenapa ya? Apakah memang gk bisa dihilangkan. Meski dgn cara operasi ? Soalnya aku tu sering ngaceng kalo lihat dia.. Tapi off lagi tiap liat jakunnya itu.. ?	0.04241
88	Kak @lucintaluna kok muka nya lama lama mirip cowok sih kak ??	-0.006
89	Bisa untuj muka juga?	-0.1607
90	Mikir banci	-0.3929
91	Bencong bercangkang	-0.3929
92	VIRAL!!!!!!!!!!!!????RAHASIA PUTIH GIGI ARTIS & SELEBGRAM GiGI ANDA KUNING? PLAK? ADA KARANG GIGINYA? BIKIN KAMU GAK PERCAYA DIRI GOSOK GIGI MU DENGAN SUPER WHITENING TEETH READYSTOCK PAKET PEMUTIH & PERAWATAN GIGI SUPER WHITENING TEETH ASLI 100RIGINAL HANYA TINGGAL 4 PAKET TERAKHIR . BERIKUT FUNGSI SUPER WHITENING TEETH ORIGINAL : 1. Memutihkan gigi dengan cepat 2. Menghilangkan noda & karang gigi 3. Menyegarkan nafas & mencegah bau mulut 4. Memperkuat gigi mencegah gigi berlubang 5. Menjaga kesehatan gigi & mencegah sekaligus mengobati gusi berdarah 6. Cocok untuk pengguna behel karena terbuat dari bahan alami tanpa bahan kimia yang membuat korosi behel untuk informasi seputar produk silahkan Follow: @pemutih_gigi79 ORDER / KONSULTASI : WA : 0821-1993-1594 FAST(RESPON)	0.06335
93	Banci banci banci banci	-0.6429
94	Assalamualaikum ga jawab banci :v	-0.3143
95	Awalnya aku emang nyangka lucinta itu perempuan. Tapi pas diperhatiin mukanya kayak banci, banyak jd yg viral viral sekarang tentang lucinta itu laki laki. Dan sekarang fix, menurut gue lucinta itu laki laki	0.01299

96	Kntl	-1.8571
97	Bencong kesasar apa yah?	-0.4107
98	Ni bencong ya	-0.5714
99	Opsi 15	0
100	Oi bencong	-0.8929
101	Dulu lu suka di tusuk pantat lu kan	-2.3393
102	Pasti keluar tai ya	-1.4107
103	Dasar bencong	-0.8929
104	Bencong Anjing	-1.7857
105	Opsi 21	0
106	Banci kontolllllllllll	-0.8929
107	Bencong / wanita berjagun wkwkwkwkwkw Uda kehabisan malu iyaaa .pansos mullu Bencong ? @lucintaluna	-0.1143
108	@abdullahfajar.r lo malah bikin dosa sendiri tolol wkwk ?	-1.1429
109	Jual jam tangan panerai authentic, tas braun buffel authentic....	-1.1429
110	@ainnisalsabila bocah bacot bgst	-2.8214
111	mana yang katanya rambutnya ditiru sama blackpink? NGAREP LU BGST	-1.1214
112	@ainnisalsabila bocah curut jembut aja blum numbuh udh bnyk bacot aja!!! ingus tu di lap anjng?	-1.1473
113	Mas fatah ngewe yuuu?	-1.4107
114	Selamat siang buat para pecinta gadget, VIVO selalu menghadirkan sesuatu yang terbaru sesuai impian dan keinginan anda semua dengan harga yang sesuai poket kalian guys? perkenalkan hanya di GALTEKINDO semua impian anda bisa jadi kenyataan seperti beli cash hp, CREDIT PAKAI DP juga Bisa CREDIT TANPA Dp, silahkan mohon budayakan membaca lokasi kita di TANGERANG CIPONDOH Syaratnya apa aja sih?? Beghhh gampang banget pokoknyaaa!! 1. Jika bossnya kerja = KTP + SLIP GAJI, 2. Jika bossnya usaha = KTP + SKU. GAMPANG BANGET KANNN TUH!!! KREDIT HP DI KAMI MAHH GAMPANG BANGETTT NGETTT NGETTTT.. PENGAJUANNYA bisa via WHATSAPP doang!! Ga perlu cape2, jauh2, buang2 bensin kemariii! TAPI jika sudah diACC, ambil hpnya kudu kemari ya boss2ku, biar pake HP BARUnya lebih afdolll gitu.. ? BENER2 TELEK DIBUKA DIDEPAN MATA BOSSNYA.. ??? note : khusus wilayah JABODETABEK yaa boss.. Bagi yg mau ajuin : bisa langsung WA ke 085977304969 Laela Terima kasih, GALTEKINDO.	-0.9413
115	Itu..blg puas ...kyk nging kuping gua	-1.25
116	Lo banci ya....	-0.5714

117	Sabunnya aman dipake coli kak???	-0.3143
118	Dijual Toyota Agya TRD 1.2 A/T 2018 Semua spek original, tidak ada yang diganti Warna silver, NO BARET KM 10rb, plat exp 2023, pajak exp 26/09/19 Plat B DKI Jakarta Ban masih bagus semua ga pernah meletus ataupun bocor, sensor parkir nyala, wiper berfungsi normal, lampu all normal. Interior bersih (pakaian wanita), AC dingin, radio bluetooth + usb bisa buat telponan juga, audio mantap, matic responsif, kaki2 bagus. Pernah dipakai di dalam kota saja, tidak ke luar kota, garansi llumar 2x, track record servis a2000 lengkap, bpkb lengkap, stnk lengkap, pajak lengkap, faktur juga lengkap. Free pelayanan emergency a2000 5 tahun HARGA OTR 135,000,000 NO MAKELAR, SAYA GA SUKA. Harga segitu biasanya masih ada overkredit atau harga khusus kredit dealer. SAYA LUNAS, GA ADA UTANG, JANGAN TEGA. Trims. WA 081380742456	-0.9434
119	Makin hari makin cantik aja kak	0
120	Sini sama abang wkwk	0
121	bangga si bencong di puja puja didunia,dia bakal nyesel akhir hidupnya bagaimana	-0.0893
122	Salfok amang jidat si om fatah	0
123	Ganteng Lucinta	0
124	hai tante ganteng banget sih	0
125	Ihhh cantik kayak babi	0.08929
126	Kok pake baju cewe bukannya pake baju cowo entar kalo pake baju cewe di kata benchong lagi	0
127	aura maco lo gakan pernah ilang samsudin.	0
128	@mingue.l@pyashinta tembak cyin tembaakkkk	0
129	Subhalallah cantiknya luar biasa	0
130	banci	1.14286
131	Tambah cantik mb lun	0
132	Aku suka lipstiknya. Kak pake lipstik apa itu. Aku suka warnanya	0
133	Nda cantik le kong lonte liar tabiar	1.08163
134	@lucintaluna yes yes yes #1 supporter kak! much love ?????????? sukses selalu cc #EnjoyYourLife	0
135	Ka luna gendutan ya @lucintaluna	0
136	Kk mkin cntikk suka liatny	0
137	Laki amat mukanya	0
138	@ainnisalsabila banyak bacod lu kaga usah ikut ikutan bocah ingusan	-0.8429
139	@ainnisalsabila dasar babi laknat	0.08929
140	@ainnisalsabila bocil sok nasehatin..... Ujung2nya minta dikepoin.. Dasar kon.tol	1.07937

141	Kontrolnya kak, silahkan diorder... Siapa tau mau balik lagi kek dulu, kalo minat DM ya kak nanti aku kasih kontrol kuda untuk mu @lucintaluna ?	1.14857
142	@ainnisalsabila BACOD	0.10714
143	KALAU YANG PROMOSI WARIA, BENCONG, TRANSGENDER KAYAK GINI... JADI RAGU MAU BELI PRODUKNYA	1.07692
144	@besty_ayu norak,, duh maaf pengikutnya ratu salome cong marah ... ?	0.07792
145	Hai Cowok	0
146	Cowok	0
147	Kq kayak laki ya ???	0
148	Wanita berktl :v	0
149	Salah satu waria tercantik di Indonesia	0
150	NGOMONGNYA MENDESAH @lucintaluna wkakakkakaka	0
151	Hai Ganteng	0
152	Bisul apa jakun ?	0
153	Ada yg nyangkut di tenggorokan	0

Based on table 4.3, total of cyberbullying text and noncyberbullying text can be illustrated by Figure 4.1.

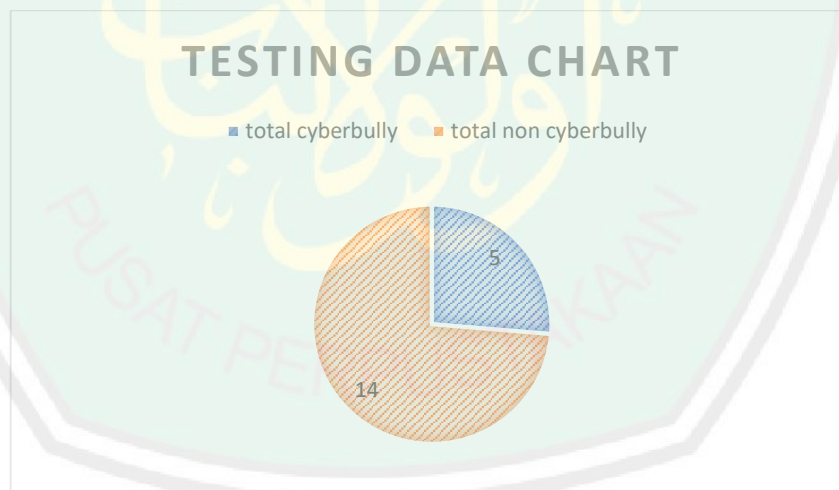


Figure 4. 1 Testing Data Chart

Comparison between real data and data output from system can be shown in table 4.4.

Table 4. 4 Result Testing Comparison

Sentences Number	Result
------------------	--------

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
1	TRUE				
2	FALSE				
3	TRUE				
4	TRUE				
5	FALSE				
6	FALSE				
7	FALSE				
8	FALSE				
9	TRUE				
10	TRUE				
11	TRUE				
12	TRUE				
13	TRUE				
14	TRUE				
15	TRUE				
16	FALSE				
17	TRUE				
18	FALSE				
19	TRUE				
20	TRUE				
21	FALSE				
22	TRUE				
23	TRUE				
24	TRUE				
25	TRUE				
26	TRUE				
27	TRUE				
28	TRUE				
29	TRUE				
30	FALSE				
31	FALSE				
32		TRUE			
33		TRUE			
34		FALSE			
35		TRUE			
36		TRUE			
37		FALSE			
38		TRUE			
39		TRUE			
40		FALSE			
41		FALSE			
42		FALSE			
43		TRUE			
44		TRUE			
45		FALSE			
46		TRUE			
47		TRUE			
48		FALSE			
49		FALSE			

50		TRUE		
51		TRUE		
52		FALSE		
53		TRUE		
54		TRUE		
55		TRUE		
56		TRUE		
57		TRUE		
58		FALSE		
59		TRUE		
60		TRUE		
61		TRUE		
62		TRUE		
63			TRUE	
64			TRUE	
65			TRUE	
66			TRUE	
67			TRUE	
68			TRUE	
69			TRUE	
70			TRUE	
71			TRUE	
72			TRUE	
73			TRUE	
74			TRUE	
75			TRUE	
76			FALSE	
77			TRUE	
78			TRUE	
79			FALSE	
80			TRUE	
81			TRUE	
82			FALSE	
83			FALSE	
84			TRUE	
85			FALSE	
86			FALSE	
87			TRUE	
88			TRUE	
89			TRUE	
90			TRUE	
91			FALSE	
92			TRUE	
93			FALSE	
94				TRUE
95				TRUE
96				FALSE
97				FALSE
98				TRUE
99				TRUE

100				FALSE	
101				TRUE	
102				TRUE	
103				FALSE	
104				FALSE	
105				TRUE	
106				FALSE	
107				TRUE	
108				TRUE	
109				TRUE	
110				TRUE	
111				TRUE	
112				TRUE	
113				TRUE	
114				TRUE	
115				TRUE	
116				FALSE	
117				TRUE	
118				TRUE	
119				TRUE	
120				TRUE	
121				FALSE	
122				TRUE	
123				TRUE	
124				TRUE	
125					FALSE
126					TRUE
127					FALSE
128					TRUE
129					TRUE
130					FALSE
131					TRUE
132					TRUE
133					FALSE
134					TRUE
135					TRUE
136					TRUE
137					TRUE
138					TRUE
139					FALSE
140					FALSE
141					FALSE
142					TRUE
143					TRUE
144					TRUE
145					TRUE
146					TRUE
147					TRUE
148					FALSE
149					TRUE

150					TRUE
151					TRUE
152					TRUE
153					TRUE

Based on table 4.4, the calculation of accuracy can be done in accordance with the design test described in section 4.1 is calculating precision and recall. The result of measure accuration based on table 4.1 can be illustrated in Figure 4.2.

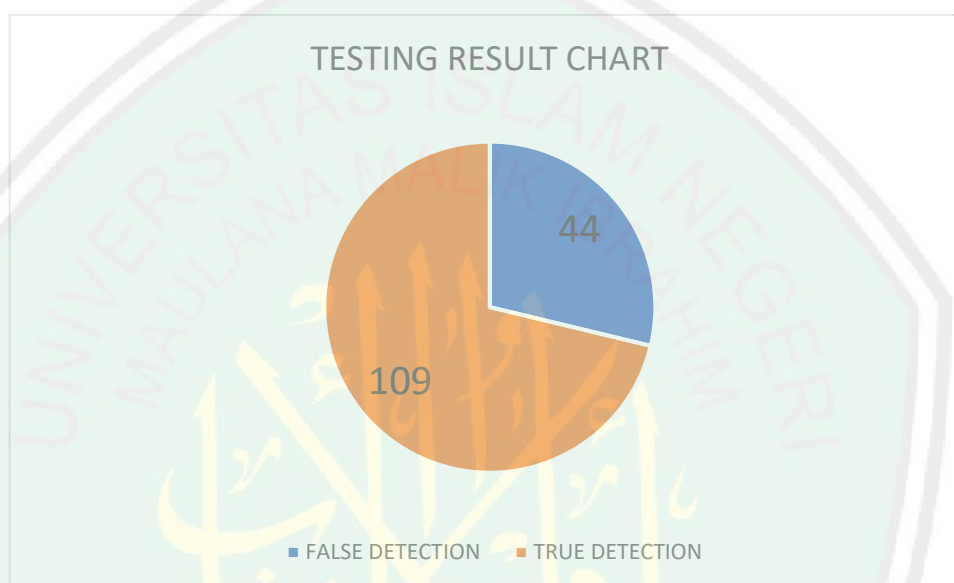


Figure 4. 2 Testing Result Chart

While, The results of these calculations can be seen in table 4.4.

Table 4. 5 Accuration Result

Aspect	Value					Average
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	
total bully	10	10	8	8	8	8.8
total non bully	21	21	23	23	21	21.8
total bully found	2	5	2	1	9	3.8
total non bully found	29	26	29	30	20	26.8
Accuratio n	61.29%	70.97%	67.74%	70.97%	82.76%	70.75%

true positive	0	3	0	0	6	1.8
true negative	19	19	21	22	18	19.8
false positive	2	2	2	1	3	2
false negative	10	7	8	8	2	7
Precision	0	0.6	0	0	0.66667	0.25333 3
Recall	0	0.3	0	0	0.75	0.21
Speed	24	24	20	32	28	25.6

The system produces accuracy with an average value of 70.75%, with a value of precision 0,25 and recall 0,21. hence, it can be said the system does not give good results. although the average accuracy is quite high, the detection of cyberbully sentences is still low. Then this indicates that the system has not distinguished all sentences correctly, there is still a sentence that should not be cyberbullying, known as cyberbullying.

2. Speed Measurement

Speed measurement can be done by calculating the average time needed to process a data divided by the amount of data. Then to calculate the time using formula 4.3 and calculate the speed using formula 4.4. As for the testing process carried out with 3 scenarios of data amount. The results of the 3 secretaries are as follows:

1. Speed Test with 40 text

First scenario is using 40 text random from comment, then the text is processed by system and get time needed to process data until return it. The results of the first scenario can be seen in table 4.5.

Table 4. 6 Speed Testing With 40 Text

Trial	Time
1	1524 ms
2	585 ms
3	598 ms
4	665 ms
5	795 ms
6	658 ms
7	415 ms
8	463 ms
9	433 ms
10	519 ms

The first scenario uses 40 data, then tries 10 times and measures the speed. Collected and calculated on average from the speed of getting the data. The average time needed for process 40 data is 665.5 ms.

2. Speed Test with 72 text

First scenario is using 72 text random from comment, then the text is processed by system and get time needed to process data until return it .The results of the first scenario can be seen in table 4.6.

Table 4. 7 Testing Speed With 72 Text

Trial	Time
1	947 ms
2	743 ms
3	1012 ms
4	829 ms
5	696 ms
6	567 ms
7	724 ms
8	558 ms
9	628 ms
10	745 ms

The first scenario uses 72 data, then tries 10 times and measures the speed. Collected and calculated on average from the speed of getting the data. The average training speed for 72 data is 744.9 ms.

3. Speed Test with 151 text

First scenario is using 40 text random from comment, then the text is processed by system and get time needed to process data until return it. The results of the first scenario can be seen in table 4.7.

Table 4. 8 Testing Speed With 151 Text

Trial	Time
1	1119 ms
2	849 ms
3	991 ms
4	980 ms
5	897 ms
6	922 ms
7	1081 ms
8	1109 ms
9	886 ms
10	991 ms

The first scenario uses 151 data, then tries 10 times and measures the speed. Collected and calculated on average from the speed of getting the data. The average training speed for 151 data is 982.5 ms.

3. Reliability Measurement

To measure reability, the testing was carried out by launching a demo program, then some people tried the program. Every one person will be given 1 minute to try. Then, in 1 min time counted how many people failed or failed to get data. Then the data is used to use the formula described in chapter 3 in the research design section.

Reliability is probability error at specified time. So to measure reliability is launch demo program to several person and try it at specified time.but in this research, reliability is probability error at sepecified person. And data of reliability measurement can be seen in table 4.8.

Table 4. 9 Data for Measurement Reliability

No	Sentences	Target	Result	Error
1	Dasar gendut	1	1	No
2	jelek lu	1	-1	Yes
3	anak haram	1	-1	Yes
4	Kalimat yg menyerupakan manusia dg hewan misal "anjing kamu"	-1	-1	No
5	Bol dobol T l k nt l	1	-1	Yes
6	Bacot anying	1	1	No
7	kampang, togam, uti, asu, jancok, dancok, kimak, keleng, tambuk, cok, asw, cvk, coeg, kontlo, bangsat, puki, tembiluk	1	1	No
8	Kalimat mengandung body shaming	-1	-1	No
9	Badan lu gemuk kayak badak/gajah. Badan lu pendek amat dah, kerdil, padahal bokap nyokap tinggi. Muka lu jerawat, di amplaz napa biar bersihan. Paha sama betis lu gede kayak tales Bogor.	1	-1	Yes
10	Banyak macam kalimat bully, saya rasa semua pilihan di atas itu akurat kalimat bullying.	-1	-1	No
11	Bego di pelihara	1	1	No
12	Kamu gendut jelek ga pantes pake baju yg lucu lucu lemak sana sini ew	1	1	No
13	Sa ae kontol	1	1	No
14	"gendut lo" "jelek aja belagu" "itu badan apa gentong" "rata amat itu badan" "tepos apanya yang mau dibanggain" "itu badan apa triplek?" "ih mukanya gradakan" "item banget lo, kaya kena areng" "udah item, gendut, jelek lagi" "suing" "heh pendek"	1	-1	Yes
15	"kok gendutan"	1	-1	Yes
16	jerawat lo itu jijik tau ga kok lo gendut jd kek badut	1	1	No
17	Apaan sih gaje	1	-1	Yes
18	Ya gini nih kalo putung rokok dikasih nyawa. Udah bego ngeyel lagi	1	1	No
19	Gendut banget lo gaada yang mau ntar	1	1	No
20	Semoga apa yang kamu lakuin berbalik ke keluarga kamu, kamu ga pantes hidup, cantik cantik kok jahat, dsb	1	-1	Yes
21	BGSD	1	1	No

22	Bangsot lu ahaha goblok nyesal ortu lu ngelahirin lu di dunia	1	-1	Yes
23	Mukanya kek tante tante, kurus bgt kek tiang berjalan gadikasih makan apa sama orang tuanya, pake kacamata soalnya buta ya, item dekil buluk, autis bgt nih anak.	1	-1	Yes
24	Banyak yg gapaham karena emang gak begitu aktif di sosmed	-1	-1	No
25	Begooo	1	-1	Yes
26	Adim Jiancooooooww cok	1	1	No

In the table 4.7, the meaning of -1 is non cyberbullying sentence, while the meaning of 1 is cyberbullying sentences. Based on data in table 4.7, it can be concluded that be possibility of an error for this program in detect the sentence cyberbullying is $11/26 = 0.423076923$ for every 26 people. A good system will have a reliability value close to 0. That means, the system will not possible get error.

4.3 Analysis and Integration

In this sub-section, we will explain the analysis of the data obtained previously. And in this sub-chapter will also be explained the integration of research with the Islamic religion. However, the analaysis data results can be explained as follows:

In measuring accuracy, the results obtained are average of accuration is 70%, precision is 21%, recall is 25%. Some data is false detection data can be shown in Table 4.10.

Table 4. 10 False Detection Data Sample

Number	Sentence
1	Muka edit 😊
2	wajanya kayak tante tante
3	mukanya boros
4	Muka nya boros bgt sii
5	gawenen ngopel coeg wkkwk
6	Cuk!! Newbie!! Bleberan!!! Dasar anak muda!! Bwahahahahaha..

7	Jangan cebok Dul, nanti pantatmu ungu
8	Tumpeh tumpeh
9	ah, masih bocah mirip babi
10	Banci kaleng ama duit mah cantik smua bisa di sulap hahaha
11	Banci halu
12	lumayan jg ni bencong
13	Ya allah si BENCONG nya nonggol lagi
14	@erih.riflan ganti otakmu bang
15	Tau gak klo ORANG BODOH ITU GAK BISA MIKIR???
16	Malunya anak2mu mas, orang tua ko sukanya nebar kebencian, syukurlah mba maia cerai kalau gak pasti malu juga...
17	Cebong tolol dan gak waras
18	Bisanya niru mulu tu cocotnya aja mrongos gitu
19	Ehh tikus curutttt, liat tu bibir nya gemeshhh
20	Goblok ee, disindir ga berasa
21	JENDERAL KARDUS, BADAN GEDE SEPERTI BAGONG, BELAGAT SONGONG OMONG KOSONG.

From table 4.10, can be explained that, some text is not found in database, and some texts false training point. But, average of false detection is because error training, and limited training data. Some example of sentence is “ cebong mana cebong, aku punya anjing”. Some wnimal word is not shown bully, so system produce false detection in training data process. Therefore, the results obtained are still threshold results that need to be re-examined using more data.

In speed measurement, the results of the time needed to process the data obtained in each scenario are 665.5 ms for 40 data, 744.9 for 72 data, and 92.5 for 151 data. This is based on the analysis conducted showing that there are differences in the speed between the amount of data even though the difference is not up to 2 times. It is likely that the time needed for search features will take a long time. This indicates that the algorithm used to search features is not good,

In reliability measurements, the results show that the system will show an error of 0.423076923 per try by 26 people. This is because, the data used in the research is very little. So, the detection text process does not show the results that are capable of arising from the data entered by each respondent. In addition, the limitations of the library used in this study are limited to the author's knowledge.

From the results obtained, it shows that the system can reduce cyberbullying crime, where this action has been banned in Islam listed in Al-Qur'an Surah Al-Hujurat verse 11:

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا يَسْخَرُ قَوْمٌ مِنْ قَوْمٍ عَسَىٰ أَنْ يَكُونُوا خَيْرًا مِنْهُمْ وَلَا نِسَاءٌ مِنْ نِسَاءٍ عَسَىٰ أَنْ يَكُنَّ خَيْرًا مِنْهُنَّ وَلَا تَلْمِزُوا أَنْفُسَكُمْ وَلَا تَنَابَزُوا بِالْأَلْقَابِ بِئْسَ الْإِسْمُ الْفُسُوقُ بَعْدَ الْإِيمَانِ وَمَنْ لَمْ يَتُبْ فَأُولَٰئِكَ هُمُ الظَّالِمُونَ

Meaning: "O you who have believed, let not a people ridicule [another] people; perhaps they may be better than them; nor let women ridicule [other] women; perhaps they may be better than them. And do not insult one another and do not call each other by [offensive] nicknames. Wretched is the name of disobedience after [one's] faith. And whoever does not repent - then it is those who are the wrongdoers." (Q.S. Al-Hujurat: 11).

In the above verse, it is explained that we as a human can not mock each other, or to bully each other. And we as human beings can not degrade others because we can be better than others. Cyberbully acts is include dzholim acts. And Rosulullah sallallaahu 'alaihi wa sallam reminded,

حِجَابُ اللَّهِ وَبَيْنَ بَيْدِهَا لَيْسَ فَإِنَّهَا ، الْمَظْلُومِ دَعْوَةٌ أَتَى.

Fear you against the prayers of those who are wronged. Because there is no veil between him and God. (Hadist Bukhari: 2448). Therefore, if the person who bullied

pray for is something bad to us, the maybe Allah will grant it, and if the person who bullied demand us in akhirat, then our charity will end.



CHAPTER V

CONCLUSIONS AND SUGESTIONS

Experiment deliver the follow results:

1. The application of SVM to detect textual cyberbully produce 70% accuracy. From 31 detection text, loss of accuration as much as 30% is due to system dictionary unknown bully word in text and false detetction in training process. In this case, 122 data is utilized for training phase.
2. The experiment to measure speed of SVM to detect textual cyberbully is by using three scenarios as follows:
 - a. Experiment by using 40 data produces the average speed of 665 ms.
 - b. Experiment by using 40 data produces the average speed of 744.9 ms.
 - c. Experiment by using 40 data produces the average speed of 151 ms.
3. The reliability is measure by testing the application directly by users to detect cyberbully based on the input form the herself. The user then verify the output of the system. Reliability is then computing by user aggrement comparement by all of the number of all users participation. Exeperiment result shows 52% reliability.

Further carried out for future research improvement as follows:

1. Increase the amount of data training and data testing because the results obtained from the system depend on how much data is obtained from the training process. it is very necessary to add data for the training and testing process.

2. Increase parameters are used to detect cyberbullying text for produce better values because in the training process system can learn a variety of sentences, so that to detect cyberbullying text will be more accurate.
3. Increase total word in dictionary are used to detect sentences. So far, the dictionary is only based on what is known by the author.



REFERENCE

- Brookover, L. (2008). State policies on school cyber bullying. Retrieved June 18, 2010, from http://www.firstamendmentcenter.org/PDF/cyber_bullying_policies.pdf.
- Bullypolice.org. (2010). Forty-three states have anti-bullying laws. Retrieved June 17, 2010, from <http://www.bullypolice.org/>.
- Dinakar K., Richard R., Henry L. (2011). Modeling the Detection of Textual Cyberbullying. Massachusetts Institute of Technology. for Safe and Responsible Internet Use.
- Haidar B, Chamoun M, Serhrouchni A. (2017). A Multilingual System for Cyberbully Detection: Arabic Content Detection using Machine Learning. *Advances in Science, Technology and Engineering Systems Journal* Vol. 2, No. 6, 275-284
- Hariani & Riadi, Imam. (2017). Detection Of Cyberbully On Social Media Using Data Mining Techniques. *International Journal of Computer Science and Information Security*,. 15. 244-250.
- Hinduja, S. & Patchin, J. W. (2018). Cyberbully Identification, Prevention, and Response. Cyberbully Research Center (Cyberbully.org).
- Hon. L.C., Varathan K.D. (2015). Cyberbully Detection on Twitter. *International Journal of Information System and Engineering* Vol. 3 (No.1), April, 2015
- Jakpat. (2018). Indonesia Social Media Trend 1st Semester 2018 [Internet]. Available in: <https://blog.jakpat.net/indonesia-social-media-trend-1st-semester-2018/>
- John A, Glendenning AC, Marchant A, Montgomery P, Stewart A, Wood S, Lloyd K, Hawton K. (2018) Self-Harm, Suicidal Behaviours, and Cyberbullying in Children and Young People: Systematic Review *J Med Internet Res* 2018;20(4):e129 DOI: 10.2196/jmir.9044
- Juvonen, Jaana, Phd & Elisheva F. Gross, Phd (2008). Extending the School Grounds?—Bullying Experiences in Cyberspace. *Journal of School Health, American School Health Association*
- Marleni Pandie, Mira & Weismann, Ivan. (2016). Pengaruh Cyberbully Di Media Sosial Terhadap Perilaku Reaktif Sebagai Pelaku Maupun Sebagai Korban Cyberbully Pada Siswa Kristen SMP Nasional Makassar. *Jurnal Jaffray*. 14. 43-62. 10.25278/JJ.v14i1.188.43-62.
- Nahar, V., Li, X., Pang, C. and Zhang, Y. (2013). Cyberbully Detection based on Text-Stream Classification. In *Proc. Eleventh Australasian Data Mining*

Conference (AusDM13) Canberra, Australia. CRPIT, 146. Christen, P., Kennedy, P., Liu, L., Ong, K.L., Stranieri, A. and Zhao, Y. Eds., ACS. 49-57

Nandhini, B. S., & Sheeba, J. I. (2015). Online Social Network Bullying Detection Using Intelligence Techniques. *Procedia Computer Science*, 45, 485–492. doi:10.1016/j.procs.2015.03.085

Nayak, Deepak & Dash, Ratnakar & Majhi, Banshidhar. (2015). Least squares SVM approach for abnormal brain detection in MRI using multiresolution analysis. 10.1109/CCCS.2015.7374140.

Nugroho, A. S., Witarto, A. B. & Handoko, D., 2003. Teori dan Aplikasinya dalam Bioinformatika. IEEE.

Pang, Chaoyi. (2013). An Effective Approach for Cyberbullying Detection. *Communications in Information Science and Management Engineering*. 3.

Patchin, Justin W. & Sameer Hinduja. (2012). *Cyberbully Prevention And Response: Expert Perspectives*. New York: Routledge

Potha N, Maragoudakis M. (2014). Cyberbully Detection using Time Series Modeling. *IEEE International Conference on Data Mining Workshop*

Purnamasari, N., Fauzi, M., Indriati, I., & Dewi, L.. (2018). Identifikasi Tweet Cyberbully pada Aplikasi Twitter menggunakan Metode Support Vector Machine (SVM) dan Information Gain (IG) sebagai Seleksi Fitur. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 11, p. 5236-5332, juli 2018. ISSN 2548-964X. Tersedia pada: <<http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/3276>>. Tanggal Akses: 24 okt. 2018

Romsaiyud, W., na Nakornphanom, K., Prasertsilp, P., Nurarak, P., & Konglerd, P. (2017). Automated Cyberbully detection using clustering appearance patterns. 2017 9th International Conference on Knowledge and Smart Technology (KST). doi:10.1109/kst.2017.7886127

Savanaraj A., Sheeba j.i., Davaneyan s.p. (2016). Automatic detection of Cyberbully from twitter. *IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS)*, ISSN: 2249- 9555 Vol.6, No.6

Singhal, P & Bansal, A. (2013). Improved textual Cyberbully detection using data mining. *International Journal of Information and Computation Technology*. ISSN 0974-2239. 3. 569-576.

Van C.H, Lefever E., Verhoeven B., Mennes J., Desmet B., Pauw G.D., Daelemans W., Hoste V. (2015). Automatic Detection and Prevention of Cyberbully. *The First International Conference on Human and Social Analytics*

Van Hee C, Jacobs G, Emmery C, Desmet B, Lefever E, Verhoeven B, et al. (2018) Automatic detection of Cyberbully in social media text. PLoS ONE 13(10): e0203794. <https://doi.org/10.1371/journal.pone.0203794>.

Willard, N. (2007). "Educator's Guide to Cyberbullying and Cyberthreats," Center.

Wirahutama, Gomar. (2011).” Ambiguisitas Peran Sebagai Variabel Moderator Terhadap Hubungan Antara Kualitas Sistem Informasi dan Kepuasan Pengguna Akhir Software Akuntansi Pada DPKAD Kota Semarang”. Jurnal Undip. Semarang: Universitas Diponegoro.

Yujun Yang, Jianping Li, & Yimei Yang. (2015). The research of the fast SVM classifier method. 2015 12th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP). doi:10.1109/iccwamtip.2015.7493959

Zheng, H., & Lu, H. (2012). A least-squares support vector machine (LS-SVM) based on fractal analysis and CIELab parameters for the detection of browning degree on mango (*Mangifera indica* L.). *Computers and Electronics in Agriculture*, 83, 47–51. doi:10.1016/j.compag.2012.01.012

Zhong L. Li H. Squicciarini A. Rajtmajer S., Griffin C., Miller D., Caragea C. (2016). Content-Driven Detection of Cyberbully on the Instagram Social Network. Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)

Zidi, S., Moulahi, T., & Alaya, B. (2018). Fault Detection in Wireless Sensor Networks Through SVM Classifier. *IEEE Sensors Journal*, 18(1), 340–347. doi:10.1109/jsen.2017.2771226