

# CLASSIFICATION OF RADICALISM SENTIMENT IN INDONESIAN DA'WAH RADICAL CONTENT THROUGH THE SOCIAL MEDIA TWITTER USING SENTIMENT ANALYSIS AND TEXT MINING METHODS IN THE COVID- 19 ERA

**Aulia Aziza; Munsyi**

Antasari Islamic State University

auliaaziza@uin-antasari.ac.id; munsyi@uin-antasari.ac.id

**Abstract:** *The Ministry of Communication and Information (Kemenkominfo) Indonesian declare the number of internet users in Indonesian country on 2019-2020 reached 196,7 million people. the total number of them, 95% of the users use the internet for accessing the social media networks such twitter, instagram, facebook, and many other things. This increase was also influenced by the pandemic period that was experienced throughout the world, especially in Indonesia, where the Government of President Jokowi set social rules in society. Almost all activities are carried out at home, especially in the field of education. this has led to a tremendous increase in internet access. the massive activity in internet access, especially social media, makes it easy for users to access information and share this information with anyone at any time. all of this can make irresponsible parties to use the opportunity to spread information that can lead to negative things and even provocations against other users on social media, especially Twitter social media. On Twitter social media there is a phenomenon known as a buzzer, namely social media users who have an impact on other users. platforms such as social media twitter are used by certain parties to form hate groups, racist communities that spread extremist ideas and incite anger or violence, promote radicalization, recruit members and create virtual organizations to carry out acts of terror organized through the virtual world. Regarding of the content in online media that has received enough attention along with various incidents that lead to intolerance in people's lives, it is a problem of radical content that gives birth to radicalism in society. in this research we try the Text Mining method using Sentiment Analysis for detecting of the direction radicalism sentiment in order to classify and detect content that contains radicalism. The classification of radical content with the method in this research can run very well, most of the content data classified according to the classification system produces an accuracy value of 66.37%.*

**Keywords:** *Islamic Radical Content; SARA; Text Mining; k-NN Algorithm;*

## **Introduction.**

The world is experiencing an extraordinary pandemic where almost all sectors of activity have been paralyzed due to the Covid-19 virus. This pandemic has made almost all activities both work-related, social relations, and education must be done at home. All of this gives rise to access to activities carried out online where the use of internet access is increasing rapidly, especially user access to social media. Social media, mediated by computer technology, interactively facilitates the creation and sharing of information, ideas, career interests, and other forms of expression through virtual communities and networks. Social media users can easily participate, share and create content anytime and anywhere. Based on statistical data, about 2.3 billion social media users in the world. Twitter has become one of the most popular media in society at the time. In social media, Twitter, there is a phenomenon known as a buzzer, namely social media users who have an impact on other users.(Saritha 2020)

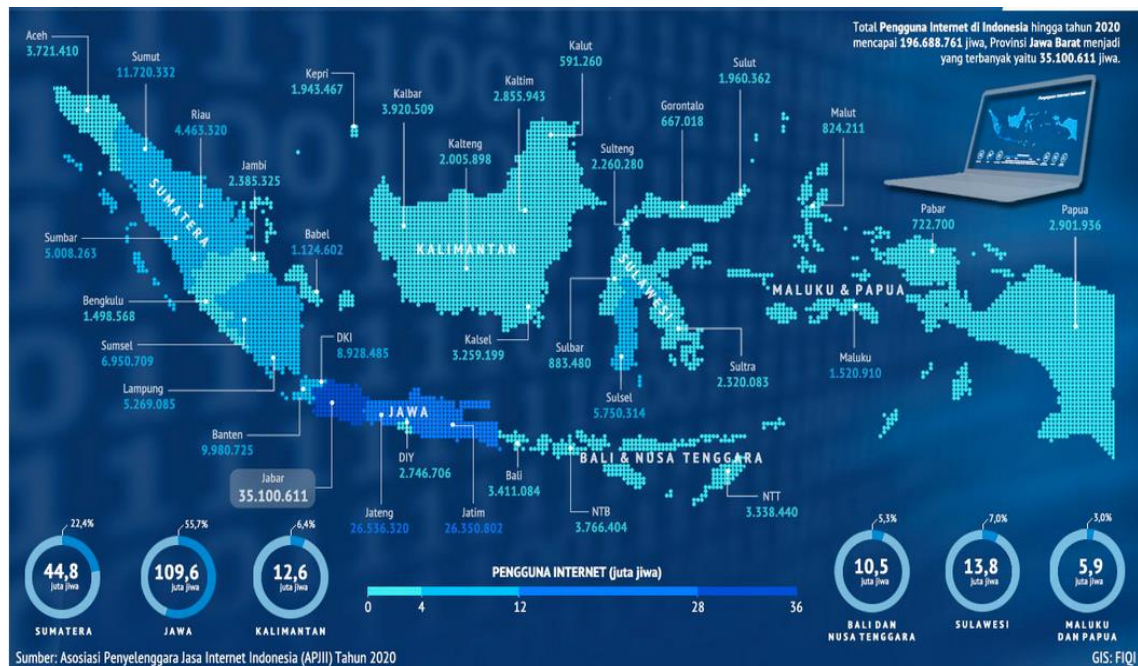


Figure 1. Radical Content Research System Design

Amount of users in the media social twitter also has the potential to make someone tend to do actions that are contrary to social norms that apply in society. The use of internet technology seems to have formed a new world society, a world that is no longer limited by territorial boundaries. Among the many uses and benefits of the development of information technology, a negative side emerges with the proliferation of crimes committed using information technology facilities (Kaplan and Haenlein 2009). This topic became of discussion by world researchers such as researchers in the fields of physiology, social science and computer science to understand and develop tools and techniques to detect it as a preventive material. Radical content in a procedural sense, namely provoking violence, spreading hatred, anti-nationalism (Setianto 2019). Indonesia's National Counterterrorism Agency (BNPT) states that radical content is a content that encourages and provokes people to commit violence in the name of religion, commit suicide and take the lives of others as jihad against people outside their group who they claim to be infidels (Agama, Negeri, and Mataram 2012).

Radicalism related to religious issues in Indonesia cannot be separated from the situation and internal powerlessness of Indonesian Muslims themselves, both nationally and internationally. Internally, polarization within the Muslim community in Indonesia occurs between various ethnic and cultural groups and organizations. In the midst of the hustle and bustle of secularization, capitalization and slander against the culture and culture of the Islamic ummah in Indonesia, both from various news reports and research reports, it shows that da'wah or disseminating information about radical ideas through social media is quite effective in attracting Muslim youth both at the level of discourse, thought and thought, as well as action. At the level of dynamics, it can be seen in online media both in discourse and thought, while at the level of action, recruitment of members of radical Islamic groups is obtained through the internet through social media and this is proven in several cases, as well as radical actions such as migrating to fight certain groups by using social media. how to carry out attacks

with sharp weapons, using firearms, even using explosives (Chaurasia and Tiwari 2013). This fact shows that something very worrying has happened to the people of Indonesia, where 87% of the population is Muslim. The lower level of media literacy for users in Indonesia is in the order of 60 out of 61 countries based on the CCSU ranking in the Press Release of the World Ranking of Nations (<http://webcapp.ccsu.edu/>), cases of spreading radical Islamic thought through social media have been reported. become a threat to the integrity and sustainability of the life of the Indonesian nation and the Unitary State of the Republic of Indonesia (NKRI).

Research shows that various platforms on the Internet (easy to publish content, allow anonymity, give exposure to millions of users and potential from on the very fast and widespread) such as YouTube (popular video sharing site), Twitter (online micro-blogging service), Facebook (popular social networking sites), online discussion forums and Instagram are easily misused for malicious purposes (Chaurasia and Tiwari 2013). Platforms such as those used to form hate groups, racist societies, spread extremist ideas, incite anger or violence, promote radicalization, recruit members and create virtual organizations and communities. This is the topic of discussion by world researchers such as researchers in the fields of physiology, social science and computer science to understand and develop tools and techniques to detect it as a preventive material. It is important for users and the public to improve media literacy skills among Muslim youth, especially among Islamic activist students, regarding radical Islamic content on social media networks.

**Method.**

In this research we using the research and development (RnD) method. Research and Development (R&D) is a research method that is carried out intentionally and systematically to improve existing products or develop a new product through testing, so that the product can be accounted for. In this research, we explain about things or theories related to the problem and the scope of the discussion are presented as a reference. The theories put forward include basic theoretical and radical definitions in terms of various views (Politics, Sociology, Culture, Economics, and Information Technology) along with theories about sentiment analysis and text mining. In addition to the term radicalism, we also know the term fundamentalist which has an interpretable meaning which sometimes means to designate a group of returnees (revivalists) in religion, where the two terms refer to an understanding where violence is a model in achieving the goals to be obtained based on theism they profess. Here are some reviews or understandings in terms of various aspects, as follows:

<b>Aspect</b>	<b>Definition</b>
Politic	Radicalism viewed from a political perspective can be symbolized as a state and power, power as a concept gives people to realize all the desires and goals of people who have that power even though the State is not the only owner of power, but the power possessed by the State is very different from the power possessed by the State. owned by other organizations.
Socilogy	Based on the understanding of conflict theory in the branch of sociology, radicalism appears as a result of the unequal distribution of authority. The unequal distribution of authority leads to the accumulation of power in one

**PROCEEDING The 3rd ICDIS 2021**  
**"Islam and Southeast Asian Communities Welfare in the COVID-19 ERA"**

---

	person, or a certain group, and with the existing authority, the group with great power will tend to use it to maintain the power they have. Or in other words, radicalism is carried out as an effort to maintain group dominance over other groups.
Culture	From this anthropological perspective, it can be understood that radicalism is a social system created by a group to defend its life from threats from other groups, as well as a means of resolving disputes that occur, either in the form of retaliation for radicalism that has been carried out, or also as a means of demanding justice which is considered to have been violated by other groups.
Economy	From the economic side, terrorist acts or radicalism are a form of resistance from the marginalized class by the oligarchs of the capitalist class (bourgeoisie) and the state. The role of the state is considered to have failed in realizing prosperity. Radicalism or terrorism is interpreted as an access to social inequality due to the exploitation of the capital of the bourgeoisie.
Religion	Understanding religion which is interpreted in a narrower sense, the meaning of jihad is misinterpreted with violent methods to find a way or impose a will in accordance with the teachings that are understood.
Information Technology	radical content in a procedural sense, namely provoking violence, spreading hatred and anti-nationalism. For BNPT, radical content is content that encourages, provokes people to commit violence in the name of religion, interprets jihad as suicide bombings and takes the lives of others, condemns people outside the group.

Text mining is the application of data mining concepts and techniques to look for patterns in the text, namely the process of analyzing text to find useful information for a particular purpose. Irregularity of the data structure in the form of text, then the text mining process requires several initial stages, which in essence is to prepare so that the text can be changed to be more structured. The steps are as follows:

1. Case Folding is the process of equating cases in a document. This is done to make the next step easier. Not all text documents are consistent in the use of capital letters. Therefore the role of Case Folding is needed in converting the entire text in the document into a specified standard form.
2. Tokenizing is the step of cutting the input string based on the words that compose it. In principle, this process is to separate every word that composes a document. In this process, punctuation numbers and other characters other than the alphabet are removed because these characters are considered as word separators (delimiters) and have no effect on text processing. Examples of this step are as follows:

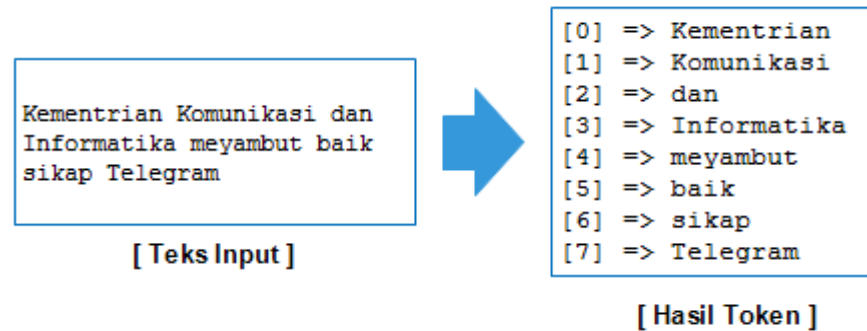


Figure 2. Tokenizing Process

3. Filtering is the step of taking important words from the token results in the previous stage. To select important words, use a stop word list. Stop word list is a list of words that cannot be used as a representative document. Examples of this step are as follows:



Figure 3. Filtering Process

4. Stemming is the process of changing word forms into basic words or the stage of finding root words or basic words from each filtered word. By doing the stemming process, each affixed word will become a basic word, thus optimizing the text mining process. The implementation of the stemming process is very diverse, depending on the language of the document. Examples of this step are as follows:

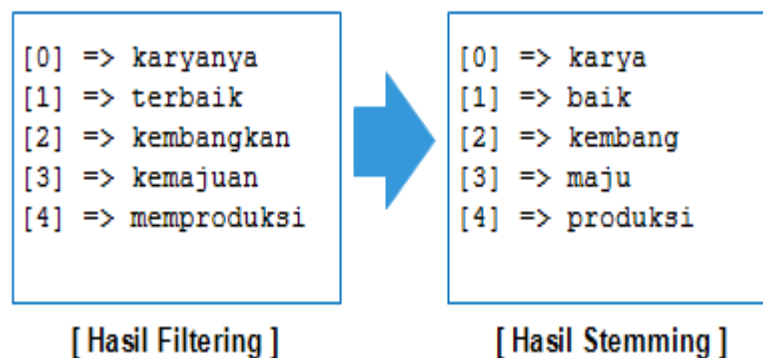


Figure 4. Stemming Process

5. Tagging is the stage of finding the initial form / root of each past word or word resulting from the stemming stage. For Indonesian text processing, tagging is not applied because Indonesian does not have a past tense. Examples of this step are as follows:

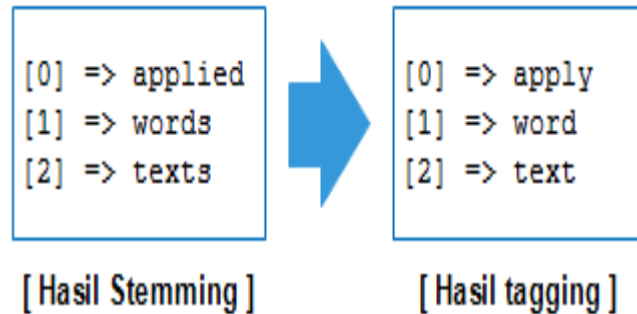
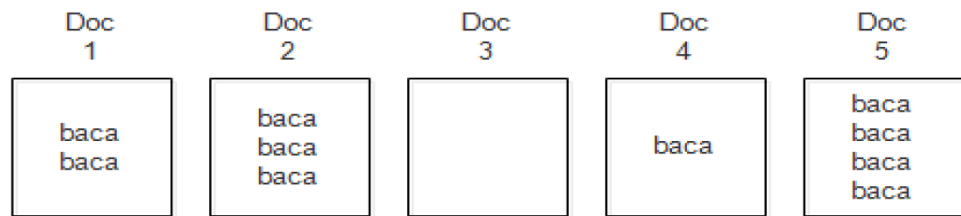


Figure 5. Tagging Process

6. Analyzing is the stage of determining how far the connection between words between existing documents is. The simplest algorithm that is usually used for scoring is TF-IDF (Term Frequency-Inversed Document Frequency). Examples of this step are as follows:



$$TFIDF_{d,t} = FREQ_{d,t} * (1 + \log \frac{N}{DFREQ_t})$$

$$TFIDF_{baca.Doc 2} = 3 * (1 + \log (\frac{5}{4}))$$

Figure 6. Analyzing Proses

7. Classification is a method in Data Mining where this method very used for solving real-world problems. As one of the most popular in the family that uses "machine learning" techniques, classification studies patterns from historical data (a collection of information such as traits, variables, features on various characteristics of previously labeled items). the goal is to place new instances (objects) (with previously unknown labels) into their respective groups or classes. The most common two-step methodologies of classification type prediction are 'model development/training' and 'model testing/deployment'. In the 'model development' phase a set of input data, including the actual class labels, is used. After a model is 'trained', it is tested against the remaining data samples for accuracy assessment and finally implemented for real use which is used to predict the classes of new data instances (where the class label is unknown). The classification method refers to the formation of data groups by applying known algorithms to the data warehouse under examination. This method is useful for business processes that require categorical information such as marketing or sales. It can use various algorithms such as closest neighbors, learning decision trees and others.

8. Sentiment Analysis is the method where in this process using natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, measure, and study affective states and subjective information. Sentiment analysis is widely applied to voice customer materials such as survey reviews and responses, online and social media, and healthcare materials for applications that range from marketing to customer service to clinical medicine. Sentiment analysis can be used to reveal public opinion on an issue, service satisfaction, policy, cyber bullying, and other analyzes based on textual data. Sentiment analysis is a computational study that discusses attitudes, judgments, opinions, emotions, and views from a set of texts that focus on extracting, identifying or finding data characteristics in the form of sentiments in the unit text using NLP (Natural Language Processing), statistics or machine methods. learning

*A. Scheme of Islamic Radical Content*

The system built is expected to help classify tweets that contain radical and non-radical content. The system that was built provides literacy to people who use social media on Twitter to avoid radical understanding and provocation of a group, both between religions, cultures and the government. The system design in this study can be seen in Figure 7.

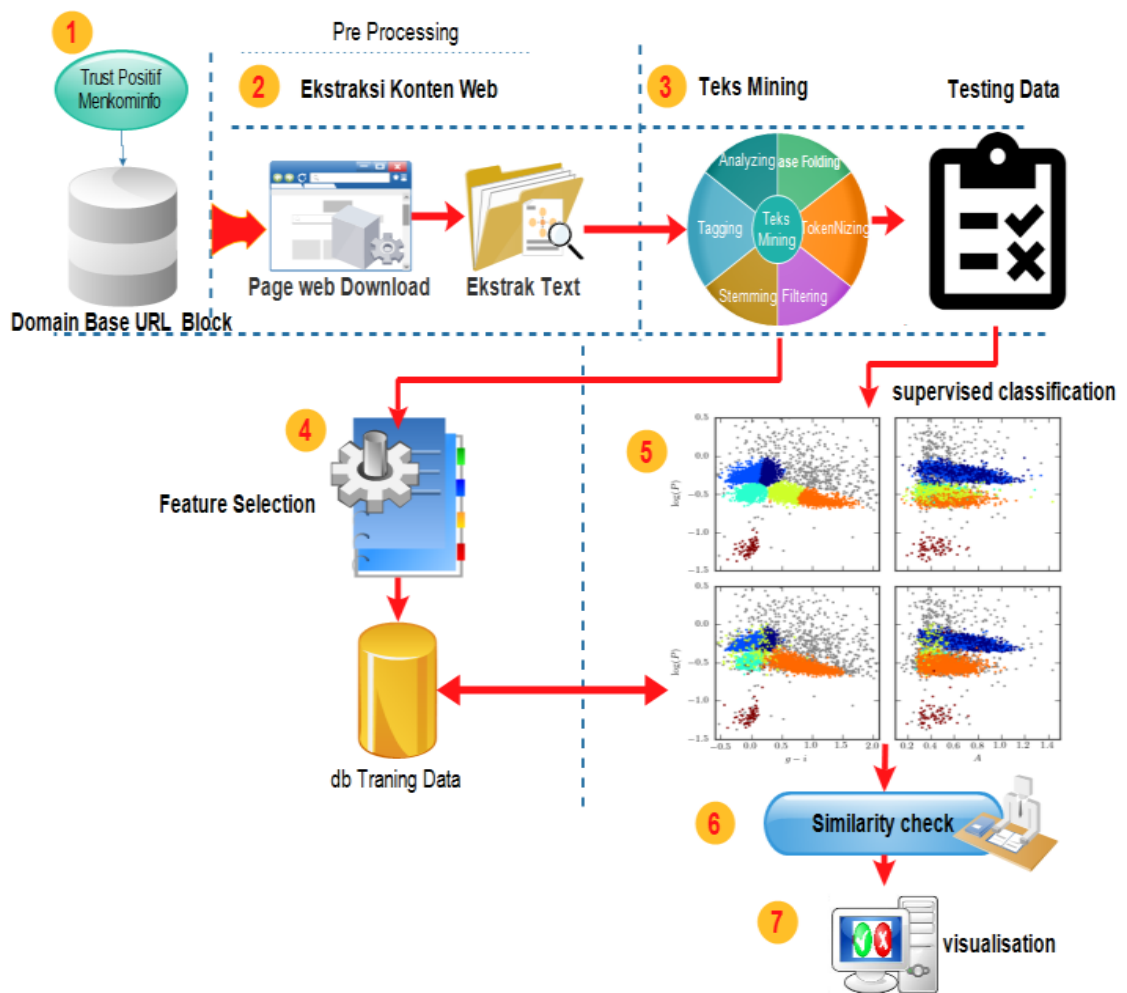


Figure 7. Radical Content Research System Design

As in Figure 7. The flow of the system starts from taking news content data from the positive internet of the Ministry of Communication and Information then the web content extraction process is carried out. Web content extraction is removing the html tag and leaving text content then the next process is Text Mining. Text mining is the process of extracting and retrieving words that represent what is in a news document, usually called keywords. The text mining process consists of several stages, namely: case folding, tokenizing, filtering, stemming, tagging and the last process of text mining is analysis. The results of the analysis are in the form of keywords (Suryandari and Lutviana 2020). This keyword became a reference for the feature selection process. The next process, these keywords are stored in the database, at the same time the features will be selected and saved back into the database. Feature selection is done to reduce the dimensional space of the data. Then the classification process. The process is the stage after the selection feature is carried out. In this classification, to determine which keywords or features correspond to the label class according to the label class that has been given by the Ministry of Communication and Information Technology. Flow of the process show in figure 8.

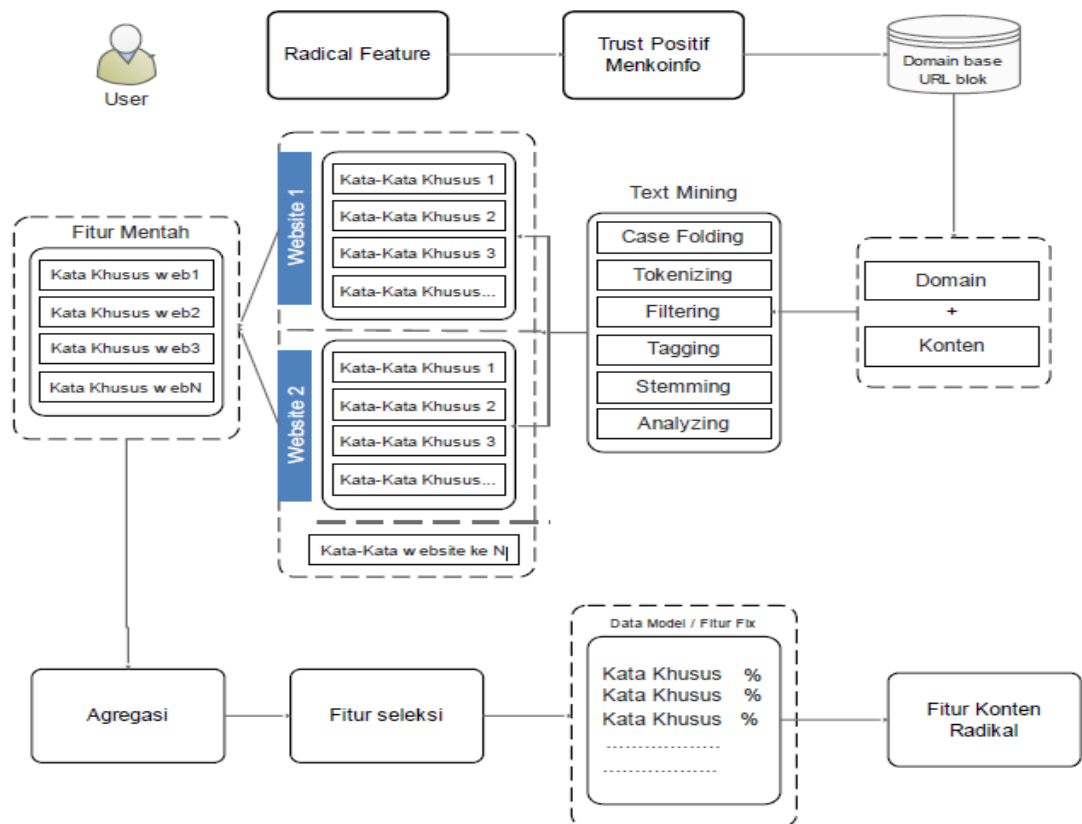


Figure 8. Design System for Learning Radical Content

The last process is a measure the level of data similarity between the data and keywords that have been stored in the database using the Euclidean distance. Then the keyword is visualized according to the criteria for the closeness of the data to the labels stored in the



database. The discussion in this stage is to get a model from radical content. The process of making this data model is carried out in several stages starting from literature study, collection of radical content data, preprocessing starting with text mining, manual class labeling, feature selection, so that a model of radical content is obtained according to the characteristics and definitions of radicals in Indonesia.



Figure 9. Process Stages in Text Mining

Our system will be collected the data, then text mining will begin. As seen as figure 9 Text Mining itself functions to look for words that represent what is in the content so that these words can be analyzed and related. Text mining is the application of data mining concepts and techniques to look for patterns in text, which is the process of analyzing text to find useful information for specific purposes. Based on the irregularity of the text data structure, the text mining process requires several initial stages, which are basically preparing so that the text can be changed to be more structured. The text mining process is carried out in 5 stages, namely Tokenizing, Filtering, Stemming, Tagging, and Analyzing. The five stages are carried out sequentially and are interconnected.

The next process is giving 4 class labels: Green, Yellow, Red, and White. Red Label is symbolized as content that has a very big influence such as (enabling readers to mobilize and raise funds, provoke to oppose the government), Here is a description of each class label content radical obtained from the Menkominfo:

Table 1. Radical Content Classification Label

Number	Classification	Description
1	Red	Calls on specific actions, mobilization of funds, and people
2	Yellow	Invitation to a certain attitude, provocation
3	Green	Spread fake news to certain groups
4	White	Normal Content

In this research the literature studies on radical content. This literature study includes the types of radical content and non-radical content, radical research in various countries, characteristics of content categorized as radical from various countries, content of radical content that can be converted into radical features of Indonesia, as well as studies on methods

**PROCEEDING The 3rd ICDIS 2021**  
**"Islam and Southeast Asian Communities Welfare in the COVID-19 ERA"**

---

of determining radical class in various countries. country. This category is then partly used as rules for radical content on the system being built. Data collection was carried out by interviewing sources related to radical Islam and search methods with search indexes for radical content, web extremes, terror and all content related to issues of SARA (ethnicity, religion, and sense of class) such as: Contents of Insult, Hate Spread, Content Websites of Religious Organizations (MUI, NU, Hidayatullah, Hisbut Tahrir, PGI, MAGABUDI, PHDI). As well as data Domain base content that is considered negative by the Government in this case the Ministry of Information and Communication (Menkominfo) on positive links trust. A total of 23 content containing SARA, 23 Content with Radical Contents, for an explicit explanation of the content of the positive trust category from the Minister of Communication and Information are shown in table 2 as follows:

Table 2. URL Positive Internet from Menkominfo

URL
1. <a href="http://ajirulfirdaus.tumblr.com">http://ajirulfirdaus.tumblr.com</a>
2. <a href="http://batalyontauhidwassunnahwaljihad.blogspot.co.id/">http://batalyontauhidwassunnahwaljihad.blogspot.co.id/</a>
3. <a href="http://anshoruttauhidwassunnahwaljihad.blogspot.co.id/">http://anshoruttauhidwassunnahwaljihad.blogspot.co.id/</a>
4. <a href="https://jalanallah.wordpress.com/">https://jalanallah.wordpress.com/</a>
5. <a href="https://religionofallah.wordpress.com/">https://religionofallah.wordpress.com/</a>
6. <a href="http://daulahislamiyyah.is-great.org/">http://daulahislamiyyah.is-great.org/</a>
7. <a href="http://ummatanwahidatan.is-great.org/">http://ummatanwahidatan.is-great.org/</a> Access
8. <a href="http://metromininews.blogspot.co.id/">http://metromininews.blogspot.co.id/</a>
9. <a href="http://al-khattab1.blogspot.co.id/">http://al-khattab1.blogspot.co.id/</a>
10. <a href="http://fadliistiqaomah.blogspot.co.id/">http://fadliistiqaomah.blogspot.co.id/</a>
11. <a href="https://daulah4islam.wordpress.com/">https://daulah4islam.wordpress.com/</a>
12. www.muhammad.com Domain Url Closed
13. <a href="https://abdulloh7.wordpress.com/">https://abdulloh7.wordpress.com/</a>
14. <a href="http://ruju-ilalhaq.blogspot.co.id/">http://ruju-ilalhaq.blogspot.co.id/</a>
15. <a href="http://fursansyahadah.blogspot.co.id/">http://fursansyahadah.blogspot.co.id/</a>
16. <a href="https://karawangbertawhid.wordpress.com/">https://karawangbertawhid.wordpress.com/</a>
17. <a href="http://terapkan-tauhid.blogspot.co.id/">http://terapkan-tauhid.blogspot.co.id/</a>
18. <a href="https://arrhaziemedia.wordpress.com/">https://arrhaziemedia.wordpress.com/</a>
19. <a href="http://syamtodaynews.xyz/">http://syamtodaynews.xyz/</a> Error Access
20. <a href="https://anshardaulahislamiyahnusantara.wordpress.com/">https://anshardaulahislamiyahnusantara.wordpress.com/</a>
21. <a href="http://jihadsabiluna-dakwah.blogspot.co.id/">http://jihadsabiluna-dakwah.blogspot.co.id/</a>
22. <a href="http://kupastajam.blogspot.co.id/">http://kupastajam.blogspot.co.id/</a>
23. <a href="https://mabesdim.wordpress.com/">https://mabesdim.wordpress.com/</a>
24. <a href="http://anshorullah.com/">http://anshorullah.com/</a>
25. <a href="http://azzam.in">http://azzam.in</a>
26. <a href="http://bahrunnaim.co">http://bahrunnaim.co</a>
27. <a href="http://dawahislamiyyah.wordpress.com">http://dawahislamiyyah.wordpress.com</a>
28. <a href="http://keabsahankhilafah.blogspot.co.id">http://keabsahankhilafah.blogspot.co.id</a>

29. <http://khilafahdaulahislamiyyah.wordpress.com>
30. <http://tapaktimba.tumblr.com>
31. <http://mahabbatiloveislam.blogspot.co.id>
32. <http://thoriquana.wordpress.com>
33. <http://tauhiddjihat.blogspot.co.id>

### **Result and Discussion.**

We implementation the system in the form of a content search engine application using the PHP programming language. It is intended that the application or system built can be accessed by any operating system and PHP is open source, meaning that anyone can develop or build applications using the programming language. This stage begins with building a search engine form to make it easier for users to search for news related content.

The user will perform a query in the form of news, or in the form of a sentence which will be queried later, the results of the query will be text mining, then compared with the words (keywords) stored in the database, then the system will check how close the word being queried is to the data in the database as a keyword. The system will issue news labels that the user searches for. In this classification process, documents that have been processed until the text frequency calculation stage will be classified based on the class that has been determined (Red, Yellow, Green, White). Class label determination is done manually with Human Brain, the process is carried out by reading all the learning content obtained from the Ministry of Communication and Informatics' positive trust data scrapping, the last stage in the system is to recognize the content that the user has entered into the system. This recognition process is carried out by classifying the testing data against the data model. The following can saw in the flow of the testing process carried out by the system built in figure 10.

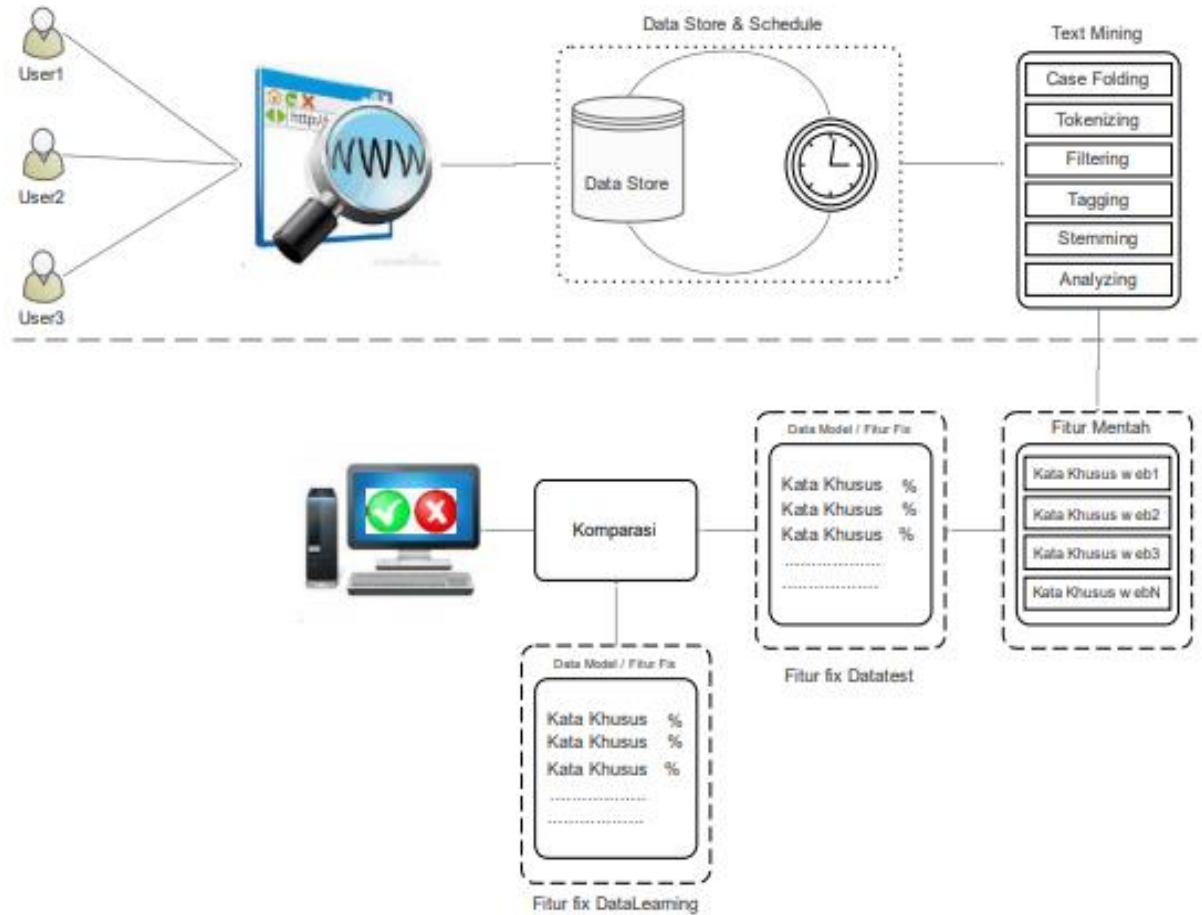


Figure 10. Design System for Training Data

The decision made by this method is that the classification process is a consideration of the test data value for all training data. The process of calculating the decision or classification in this method is to calculate the value of the proximity or distance of the test data to the overall training data. The calculation of the distance in this study was carried out by calculating the value of the euclidean distance, then this method voting for the distance values that have been sorted from the smallest to the largest value. The number of votes is a characteristic of the KNN which is denoted by the variable  $k$ . because of the nature of this algorithm based on the majority of votes. Then the  $k$  values are generally odd like 1,3,5,7, etc. From the voting results, the classification of the KNN method will be obtained on the test data against the training data. At this stage each test data will get a label that has been classified by KNN. The next step the system will try to search for data with keywords in the database, calculate the distance for each word using the euclidean distance, the result is a label (Red, Yellow, Green and White).

In this testing for the method used refers to the measurement by measuring the level of accuracy of the kNearest Neighbor algorithm based on a radical dataset which is divided into several variables or decision-making attributes. The system was built using a Web Scrapping application. the following is a flow chart of the tests carried out. In this study, there are experimental parameters including Query Results to the System, measurement of classification

accuracy using the Confolution Matrix. Classification processing time performance. This research was conducted at the Islamic Communication and Broadcasting Lab, Faculty of Da'wah and Communication Sciences, Antasari State Islamic University, Banjarmasin. For all the equipment provided by the Research Team. System testing carried out by computers with certain specifications Computer specifications for the testing process refer to the computational load applied by the developed system. The greater the value of the CPU or hardware component, the faster the computing process is carried out. The experimental results in this study consisted of 3 experiments, including: testing the query system, measuring the accuracy of the classification algorithm applied to the system. As well as testing the performance of the system built by measuring processing time and system memory usage. Systemically, the process of retrieving news content from social media twitter portals is carried out using scripts. However, for this experiment, the author takes the news content directly from the scrapping database of twitter content retrieval.

+ Options			id_content	content_web	class
<input type="checkbox"/>	Edit	Copy	1	Anshar Daulah Islamiyah Nusantara – Perjalanan hij...	MERAH
<input type="checkbox"/>	Edit	Copy	2	Kebangkitan Semu Nasionalisme Versus Kebangkitan S...	MERAH
<input type="checkbox"/>	Edit	Copy	3	Polisi Pembohong!! Tak Ada Satu-pun Bekas Luka Tem...	MERAH
<input type="checkbox"/>	Edit	Copy	4	Aksi Brutal Densus 88 Mulai Mendapat Balasan?	PUTIH
<input type="checkbox"/>	Edit	Copy	5	Dari Jihad Jadi Fasad	PUTIH
<input type="checkbox"/>	Edit	Copy	6	Daulah Islamiyyah adalah Imamah yang Hak	PUTIH

Figure 11. Query Result

The data will be directly tested by the system. The experiment started from testing the system to do text mining. Text mining is done by the system to retrieve important words in the news content. The text mining process is not much different from the training process as described in the previous chapter. The difference is in the training process after text mining is followed by a feature selection process, but the data testing process is not carried out by a feature selection process but is directly stored in the database for classification with the training database.

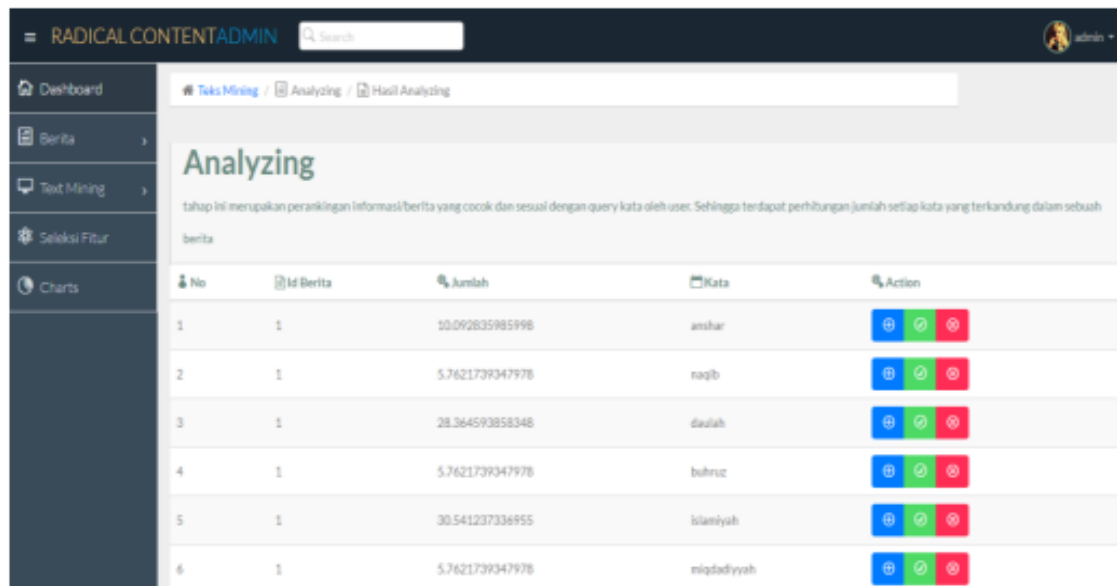


Figure 12. Analysis Result

For work that is measured in this application is the level of accuracy of the algorithm in classifying documents correctly, into the four categories mentioned above. The following is a calculation to find data accuracy.

$$\begin{aligned}
 \text{Akurasi Kebenaran}(\%) &= \frac{\sum \text{koreksi benar}}{\sum \text{data uji}} \\
 &= \frac{77}{116} * 100 = 66.37 \%
 \end{aligned}$$

From the equation, the accuracy results are 66.37%. Testing using cross validation. In addition, the accuracy to test the correctness of the results of the manual calculations above. The results of the classification above are then analyzed for the results of classification accuracy. in this study using the calculation of cofution matrix. Cofution matrix calculation is carried out on True Positive items or variables with the following equation:

$$\begin{aligned}
 \text{True Positif}(TP) &= \frac{TP}{TP+FN} \\
 \text{akurasi} &= \frac{TP+TN}{TP+TN+FP+FN} = \frac{TP+TN}{N}
 \end{aligned}$$

The test results are used to get the maximum k value in order to produce optimal success. In this study, k values of 1 to 30 were used but not all k values gave significant results, so only significant k values were selected. As shown in the following table.

Tabel 3. Optimal Result

k	Accuracy %	Error Classification
1	50.17	49.83
2	52.59	47.41
3	58.62	41.43
4	62.07	37.93
5	58.62	41.38

6	66.37	33.63
7	66.37	33.63
8	66.37	33.63
9	66.37	33.63

From table 3, it can be seen that the optimal and stable k value is obtained at k 7 and so on, so that the optimal k is obtained to get the best accuracy value, namely k = 7. The following is the k-optimal graph of this research.

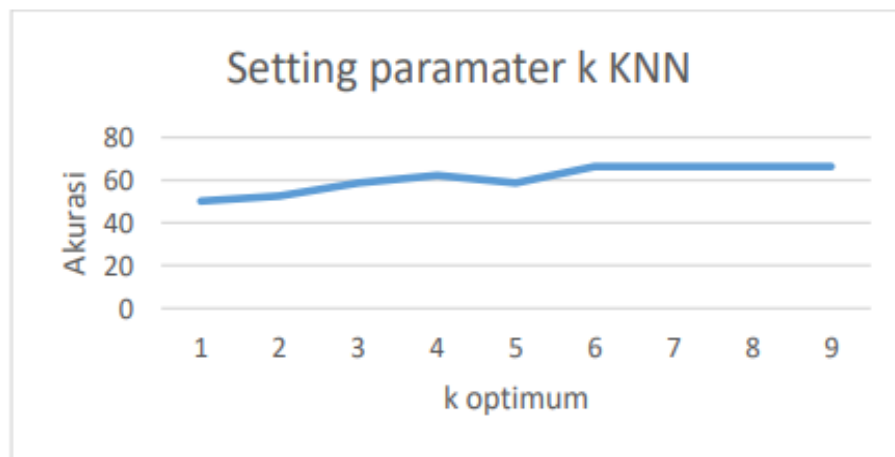


Figure 13. K-Optimal

Figure 6 shows the optimal results obtained with a value of  $k = 7$  this is due to the spread in the value of the large distance range between classes and others. The number of  $k$  determines the amount of data that thinks about the closest distance then the mode is calculated for predictions in the most classification class. After knowing the performance of the classification, then we try to measure the precision and recall values for each class to find out how close the relationship between classes from each class is. The following are the results of the precision between classes.

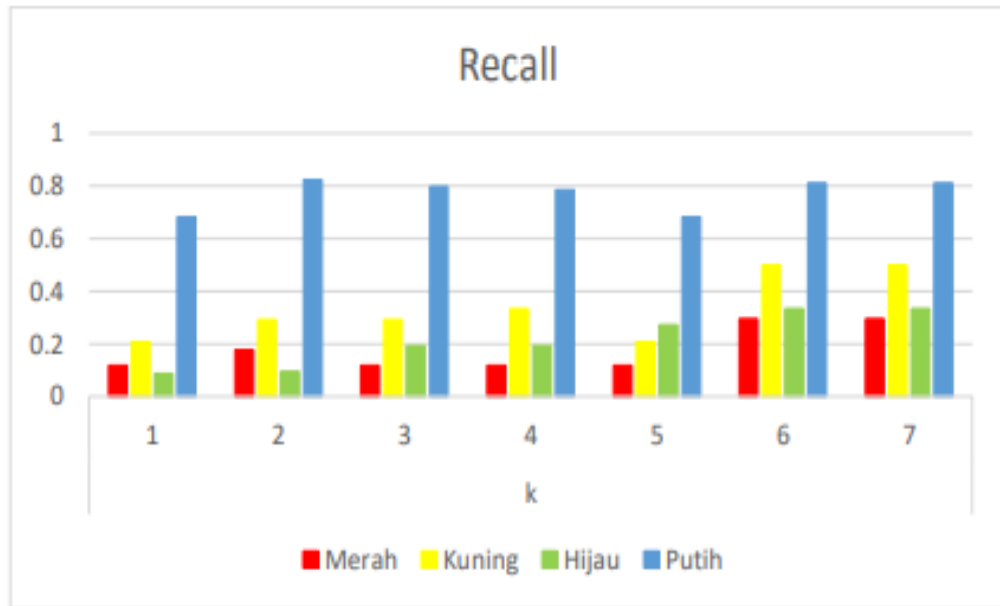


Figure 14. Recall data in every class labelling

When system compared on the values of the two tables above, it can be seen that the "White" class  $k_1 = 0.689$ ,  $k_2 = 0.824$ ,  $k_3 = 0.797$ ,  $k_4 = 0.784$ ,  $k_5 = 0.683$ ,  $k_6 = 0.811$  and  $k_7 = 0.811$  with an average increase of 0.771. The highest score is obtained at  $k_2$ , when compared to all classes. The "White" class has the highest precision and recall value in all  $k$  values.

### Conclusion.

After several experiments, conclusions can be drawn from this research, as follows:

1. Classification of radical content using the K-Nearest Neighbour algorithm carried out in this study can work well, most of the content data is classified according to manual classification with an accuracy value of 66.37%.
2. The percentage of accuracy is obtained at the optimal  $k$  value = 7 by calculating the accuracy value using cross validation, the value is 66.37% with an error rate of 33.63% of the total content of 116 content. With the comparison table (P) Precision (R) Recall, and (F) F-Measure, the highest class value is obtained in the "white" class.
3. The system that was built requires a processing time of 0.704 second and the largest memory usage is 884656 bytes, this indicates that the k-NN classification algorithm in memory usage is quite high in each classification processing.

For enhanced the system development, we recommended to add the Indonesian synonym identification function, so that if the entered keywords have synonyms with the word in the database it is not considered a new word so that it helps in the classification process, and the determination of the radical classification of themes for each label class is carried out by experts or people who are experts and responsible for determining the radical content. Maybe the next research can add or be combined with



algorithms such as Naïve Bayes Classification (NBC), Support Vector Machine (SVM), Decision Tree (Decision Tree), Neural Network (NN) to improve accuracy results. Apart from that, from the future system, this system can be developed further towards the Proxy, Mobile platform and so on.

**References**

- Agama, Institut, Islam Negeri, and Iain Mataram. 2012. "Islam Fundamentalism Sebagai Fenomena Politik" 15: 264–74.
- Chaurasia, Nisha, and Akhilesh Tiwari. 2013. "Efficient Algorithm for Destabilization of Terrorist Networks." *International Journal of Information Technology and Computer Science* 5 (12): 21–30. <https://doi.org/10.5815/ijitcs.2013.12.03>.
- Kaplan, Andreas M., and Michael Haenlein. 2009. "The Fairyland of Second Life: Virtual Social Worlds and How to Use Them." *Business Horizons* 52 (6): 563–72. <https://doi.org/10.1016/j.bushor.2009.07.002>.
- Saritha, K. 2020. "Block Chain Authentication Using Elliptic Curve Digital Signature Algorithm" 3 (2): 71–75.
- Setianto, Widodo Agus. 2019. "Literasi Konten Radikal Di Media Online." *Jurnal Ilmu Komunikasi* 16 (1): 75. <https://doi.org/10.31315/jik.v16i1.2684>.
- Suryandari, Nikmah, and Vivi Lutviana. 2020. "Impression Management of Buzzer in Social Media Twitter." *Jurnal Studi Komunikasi (Indonesian Journal of Communications Studies)*. <https://doi.org/10.25139/jsk.v4i3.2665>.