EasyChair Preprint

№ 10877

Twitter Sentiment Analysis on Long Acting Contraceptive Methods in Indonesia With Machine Learning Approach

Nurul Sari, Ahmad Munir, Muhammad Ramli and Madris Iskandar

September 10, 2023

Twitter Sentiment Analysis On Long-Acting Contraceptive Methods In Indonesia With Machine Learning Approach

Nurul Puspita Sari 1[1], Ahmad Munir 2[2], M.Ramli AT 3[3], Madris Iskandar 4[4]

[1]Nurul Puspita Sari, Master's Degree Program in Sociology Population Concentration, Department of Sociology, Hasanuddin University, Makassar, South Sulawesi, Indonesia
[2]Ahmad Munir, Department of Agricultural Engineering Studies, Faculty of Agriculture, Hasanuddin University, Makassar, South Sulawesi, Indonesia
[3] M.Ramli AT, Department of Sociology, Faculty of Social and Political Sciences, Hasanuddin University, Makassar, South Sulawesi, Indonesia
[4] Madris Iskandar, Department of Economics, Faculty of Economics and Business, Hasanuddin University, Makassar, South Sulawesi, Indonesia
nurul.ps@bps.go.id

Abstract. High population growth rates still need to be solved in almost all parts of the world. Over the past decade, Indonesia has been ranked fourth in the world regarding population. The family planning program has produced positive results. However, changes in the organizational structure of the institution and changes in local government commitment have led to high disparities in the family planning program. Long-term contraception is considered adequate as the primary need to reduce the rate of population increase. However, the pattern of contraceptive selection in couples of childbearing age in Indonesia is still dominated by non-long-term contraceptive methods such as pills and injections. In contrast, the rate of long-term contraceptive use continues to decline yearly. This study aims to see how the Indonesian people respond to long-term contraceptive products from January 1, 2020, to November 11, 2022. The research used data scraping techniques through the Twitter social media application by grouping sentiment based on negative, positive, and neutral classifications. Each sentiment was analyzed with a word cloud using keywords related to long-term contraceptive methods. Furthermore, classification evaluation is carried on by examining the accuracy of machine learning classification algorithms, specifically naive bayes and random forest.The results stated that the Indonesian people's response to long-term contraceptive methods is still negative, which means that long-term contraceptive products still need to be in demand by the Indonesian people. Based on the accuracy results, the random forest algorithm is very good at classifying tweets, which is 99,33% compared to the naïve bayes algorithm.

Keywords: tweet, Twitter, contraception, LARC, naive bayes, random forest, machine learning

1. Introduction

High population growth rates still need to be solved in almost all parts of the world. Over the past decade, Indonesia has been ranked fourth in the world regarding population. The 2020 Population Census data recorded an increase in the number of people in Indonesia by 32.56 million compared to a decade earlier to 270.20 million. Despite this, the Indonesian government has made exemplary achievements in reducing

the fertility rate to 1.28 in 2020 [1]. This suggests that the family planning program, which has been advocated since the 1980s, has produced beneficial effects. Indonesia's development has improved the quality of people's lives and social situations. Changes in the governance of the family management programs as the reform era began have reduced the government's commitment to the population control strategy[2]. Although the national accomplishment of the fertility transitions is clear, differences in features between areas, as well as the decentralization strategies of each local government, cause significant fertility differences within Indonesian provinces.. [3].

Family planning promotes self-actualization, empowerment, health, and well-being, reducing maternal and infant deaths by preventing unintended pregnancy and unsafe abortion [4]. Effective contraceptive methods are a cornerstone of unintended pregnancy prevention [5]. The following are the consequences of ineffective and useless contraception: unplanned pregnancies, which are associated with higher abortion rates [6], poorer postpartum outcomes, an incidence of postpartum maternal depression, premature and low birth weight babies, stunting, and reduced breastfeeding opportunities[7]. Long-acting reversible contraceptives (LARCs) have the advantage of having a high effectiveness rate of almost 99% and a long duration of use so that users do not need to go back and forth to insert the device or take medicine regularly. In addition, the contraceptive device in the form of an implant only consists of the hormone progesterone, which has no effect on other hormones, and the installation of the IUD, which is done by professionals, makes users feel more comfortable[8] Although the LARC technique has a higher level of effectiveness, ease of use, and satisfaction than non-long-term contraceptive methods (SARCs), the use of oral contraceptive pills is still the most widely used method in the community, while the LARC method is still underutilized. This could be attributed to a misunderstanding of the risks and benefits of LARCs, differences in access and attitudes toward them, fear of long-acting contraceptive side effects, community myths, socio-cultural norms, and religious beliefs that make the use of long-acting contraceptives less common in developing countries[9]. The percentage of LARC use in Indonesia is still low, which is only 13.4% (2017), 23.1% (2018), 23.5% (2019), 24.5% (2020), 25.93% (2021), and 22.6% (2022) from the target of 28.39% in 2024[10].

Social media plays an essential role in the lives of people of all ages in today's digital age. People of all ages use social media to communicate as well as to gather information [11]. According to the most recent data, 38.5% of Twitter users are of productive age, with a range of 25-34 years. Furthermore, there are six nations with more than 10 million Twitter users, with Indonesia ranking sixth in the world in terms of Twitter users. [12] Based on these facts, we may acquire a unique insight into the discussion that represents the population's social standards, behavioral goals, and attitudes via the Twitter application [13]. A qualitative study found that social media is frequently the primary and most trustworthy source of contraceptive information. [14]. Contraceptive use among women is primarily influenced by their own ideas as well as recommendations from friends and family regarding which contraceptive to use [15]. According to one study, after noticing brand references in Tweets, more than half of Twitter users reported taking action (searching for, reading about, and purchasing). [16]. A meta-analysis study on psychological therapies based on social networking sites

used by patients in the healthcare field revealed that such interventions are beneficial in promoting health behaviour change[11]. Over the past thirteen years, Twitter has had the potential to be a significant source of data on the subject of contraceptive discourse and individualised use among its users. Tweets can provide us with a better understanding of the demographics of the population. Recognising the impact social media has on people's lives and, perhaps, when considering the use of contraceptive methods, these and other social media platforms allow clinicians and researchers to obtain and potentially disseminate correct information about contraceptive alternatives. [17]. Through the use of social media as a data source, we can study people's views on contraceptive practices and obtain data on people's thoughts and the stigma that has developed over the years regarding the use of contraceptive methods. Traditionally, population thinking has made it difficult to engage in research related to the topic of stigma. The opinions and behavioral attitudes of adolescents and young adults, who are the productive age population, have been easier to counsel through social media because they can post, search, and even be influenced by social media content related to contraception and reproductive health [18]. Twitter sentiments can be used as a proxy for the mindset of patients considering contraception options without the influences or biases found in more traditional research methods such as surveys and in-person interviews[19]. Based on this description, this study was carried out to determine the long-term contraceptive acceptability of techniques and the effectiveness of LARC supplied by evaluating sentiment reported via the Twitter application in Indonesian. This study is aimed at providing an overview of the general public's acceptance of LARC services in Indonesia. So that it can be utilized as a government evaluation of long-term contraceptive products that have been introduced.

2. Research Method

This research is conducted in stages and structured utilizing the machine learning (ML) algorithm approach at each level, allowing it to drive the process that is deemed most appropriate. To execute study findings, this research approach employs methodologies utilized in data formulation, analysis, and collection. The dataset is derived from the Twitter data scraping technique, and it contains queries related to the use of long-term contraceptive methods via keywords (family planning, IUD, implant, tubectomy, vasectomy) as a classification of long-term contraceptive methods from January 1, 2020 to November 11, 2022. The data obtained consisted of 4,196 data points about long-term contraception.

To obtain Twitter tweets, perform the following steps:

1. Enable the rtweet and dplyr libraries;
2. Create Twitter developer originality credentials to obtain the Twitter developer API;
3. Scraping Twitter using the API with the required tweet language being Bahasa Indonesia, lang = "id".

The research design scheme is presented in the figure below:

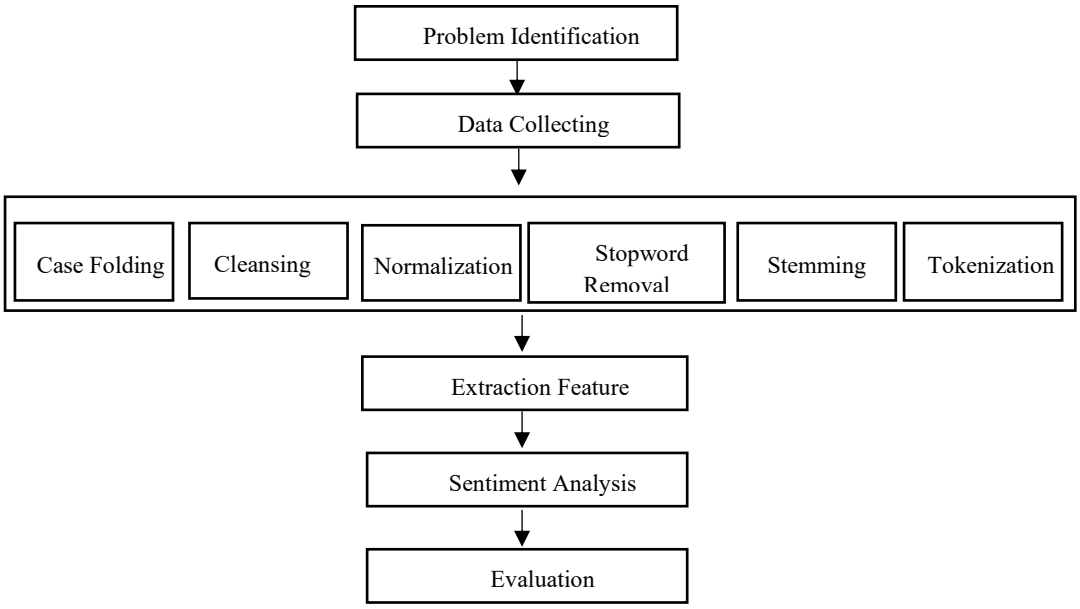


Fig. 1. Research Design

After the data was collected, the data cleaning process was carried out with the following steps:

1. Case Folding
   Case folding attempts to transform all letter forms in a text to lowercase letters. Case folding is followed by eliminating characters from the sentence.[20]
2. Cleansing
   Pre-processing is a critical stage in sentiment analysis, yet it receives little attention in the literature or models. The data obtained from various sources may contain redundancy and duplicates; any occurrence of redundancy in the datasets must be detected. In sentiment analysis, data cleaning removes redundant and inaccurate data values that are intended for analysis.[21].
3. Normalization
   Text Normalization is a data mining operation in which text is cleaned of unwanted tags and symbols. The practice of replacing HTML addresses and URLs with blanks and detecting duplicate, corrupt, and erroneous entries is known as normalization. The cleaning process continues by replacing emojis in sentences, dot characters in sentences with blanks, comma characters in phrases with blanks, characters and numerals in sentences with blanks, deleting mentions and hashtags, and tidying up sentences. [22].
4. Stopword Removal
   Stopword removal is the process of deleting unnecessary words from a description by examining the words in the parsed description to see if they are on the stoplist. If the terms are on the stop list, they will be eliminated from the description, leaving just the crucial or keyword words. Either a stoplist algorithm (which removes less

important words) or a wordlist algorithm (which keeps significant words). Stop listing is the process of deleting frequently used terms from papers, such as and, or, not, and so on. [23]

5. Stemming
   Stemming is one of the pre-processing steps used in many research to enhance sentiment classification performance. The stemming process is carried out to remove every affix word in the word, either at the beginning of the word or the end of the word, by paying attention to common prefixes and suffixes that can be found in the word and removing duplicate text again in the position of clean text.[24]
6. Tokenization
   Tokenization helps to divide the textual information into individual words. The tokenization stage is carried out to remove unnecessary punctuation and cut the text into words, symbols, characters, or punctuation marks so that it becomes a token that can be analyzed[25].
7. Extraction Feature
   When utilizing machine learning approaches to extract sentiments from subjective text, feature extraction plays an important role. A word vector will be determined from the tokenized dataset throughout the feature extraction procedure. A word vector is an algorithm that converts words into vectors. This vector will be employed in a variety of natural language processing tasks[26]

In this study, the features extracted from text sentences using feature extraction approaches are trained and assessed using the Naive Bayes Classifier (NBC) and Random Forest Classifier.

2.1. Naive Bayes Classifier (NBC)

The Naive Bayes Classifier (NBC) approach for text classification uses the properties of words in a document as the basis for classification. The classification method use the probability theory created by British scientist Thomas Bayes to forecast future probabilities based on existing experience. [27]. This algorithm works by determining the probability of a character that belongs to a specific class (sentiment) or posterior, which is then multiplied by the occurrence of the text in each tweet in a particular class (sentiment) or likelihood[28]. The mathematical form of the algorithm This algorithm is written as follows:

$$P(H|X) = \frac{P(X|H)xP(H)}{P(X)}$$

Description:

X : Unclassified data the sentiment

H : Hypothesis that X falls into a certain sentiment classification (positive, neutral, negative)

$P(H|X)$: Probability of hypothesis H based on condition X

$P(X|H)$: Probability of X based on hypothesis H

$P(H)$: Hypothesis probability H

2.2. Random Forest

Random Forest is one of the machine learning (ML) algorithms introduced by Leo Breiman and Adele Cutler. Random Forest combines the outputs of multiple decision trees to achieve a single result. This method is used for classification and regression with a decision tree aggregation tool consisting of internal, root, and leaf nodes. Random forests use recursive binary splitting to reach the final nodes in a tree structure based on classification and regression trees. Random forest is a combined decision tree algorithm that groups "weak learners" together to form "strong learners" through ensemble techniques [29].

Breiman developed the random forest technique, which has the advantages of creating relatively low errors, strong classification performance, effectively processing large-dimensional training data, and being an effective method for predicting datasets with missing data.[30].

The Random Forest Algorithm in this study is applied with the following steps:

1. Determining the number of trees k from as many features as m, with $k < m$;
2. Determining the number of k-folds of 10 for randomization and randomization;
3. In each tree that is formed, a subset of predictors (p) with $m < p$;
4. The next step is looping back as many as k trees;
5. The predicted fit classification is obtained by aggregating the most votes from all trees formed.

2.3. Model Accuracy

Model accuracy is a measure of the performance of the algorithm used in the study. This model accuracy is the basis for making the best algorithmic decision in the analysis. The confusion matrix is used in machine learning classification methods to assess the effectiveness of a classification model or classifier. In essence, the confusion matrix comprises information that compares the system's classification results to the expected classification results.

Table 1. Configuration of Confusion Matrix

Prediction	Reality	
	Positive	Negative
Positive	TP	FN
Negative	FN	TN

Description:
TP : Actually Positive (True Positive)
TN : Actually Negative (True Negatives)
FP : Not Actually Positive (False Positive)
FN : Not Actually Negative (False Negatives)

Mathematically, the accuracy measure model is written as follows:

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Classification models expect a high accuracy value. Accuracy category models can be classified as the following:

Table 2. Model accuracy category

Accuracy Value	Category
90% - 100%	Very good
80% - 90%	Good
70% - 80%	Fair
60% - 70%	Deficient
50% - 60%	Fail

Precision is defined as the fraction of true positive examples among those labeled as positive by the model. Recall, also known as sensitivity, is the percentage of positive examples among all positive examples. The F1-score is the harmonic mean of the approaches and includes the model's recall and precision. Accuracy is the most extensively used way for validating models in machine learning's numerous classification methods. [31]. This study uses the accuracy measure to see which method is better at predicting netizens' sentiment analysis on long-term contraception.

2.4. Latent Dirichlet Allocation Method (LDA)

Latent Dirichlet Allocation, or LDA, is an algorithm for topic detection through probabilistic modeling of a set of data. This method is used to find topics that are contained in Twitter tweets. The purpose of applying the LDA method is to find out what topics are being discussed by internet citizens (netizens) regarding the use of long-term contraceptive methods. This is motivated by the frequency with which Indonesian people express all their conditions on social media. Form of analysis through word clouds, a set of words in a text that illustrate the frequency of use. If the frequency of use of a particular word is large or dominant, then the size of the word in the word cloud is also large. The larger the size of a word, the more often it is used in expressions on social media.

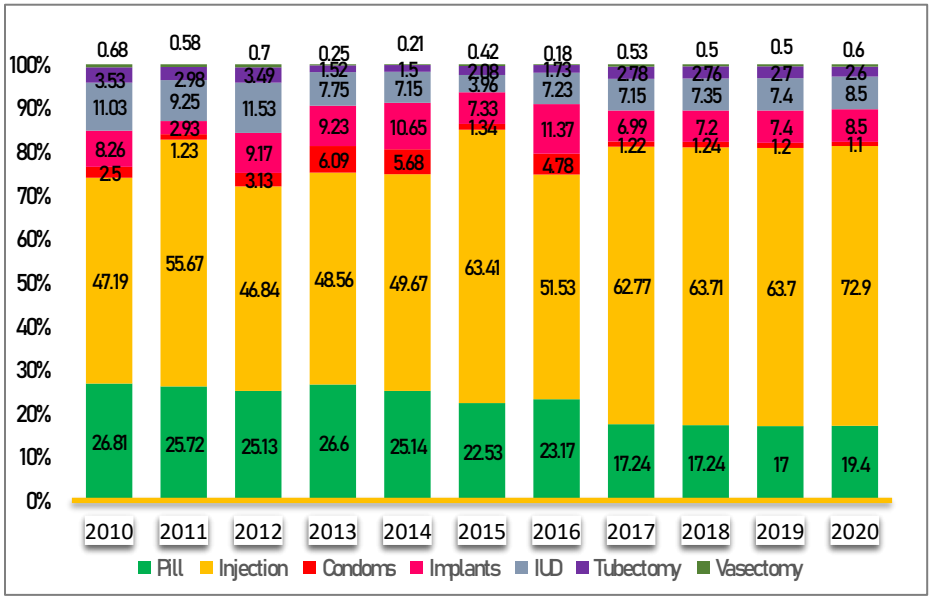
In this research, the Dirichlet Allocation method was limited to describing the words often appearing in Twitter tweets. The frequency of frequently discussed words is represented in a word cloud based on netizen sentiment. The size of the words in the word cloud was taken from the top 20 words frequently used by netizens in response to the topic of long-term contraceptive use. The frequently occurring words were analyzed to illustrate how positive and negative sentiments were formed.

3. Results

3.1. Data Exploration

.

Over the past decade, non-long-acting methods have gradually become Indonesia's primary choice of contraceptive users [32]. This is also reflected in the Indonesian Health Profile 2010–2020 data, where almost all provinces experienced a shift in the use of long-acting contraceptive methods to non-long-acting contraceptive methods such as pills and injectables[33]. Figure 2 shows that the use of contraceptive methods in Indonesia from 2010 to 2020 tends to be non-long-term methods (Non-LARC), especially the injectable method, which has a percentage of more than 40 percent, and the pill method, which has a percentage of more than 15 percent. In LARC, the use of the IUD method tends to decrease from 11 percent in 2010 to 8 percent in 2020. The use of implantable contraceptives also experienced unstable fluctuations in 2016–2020, experiencing a significant decline from 11.37 percent to 8.5 percent. Tubectomy (female method) decreased from 2010 by 3.53 percent, while vasectomy (male method) stagnated at 0.6 percent for ten years. This shows that the knowledge of family planning acceptors regarding more appropriate and effective long-term contraceptive methods is still minimal.



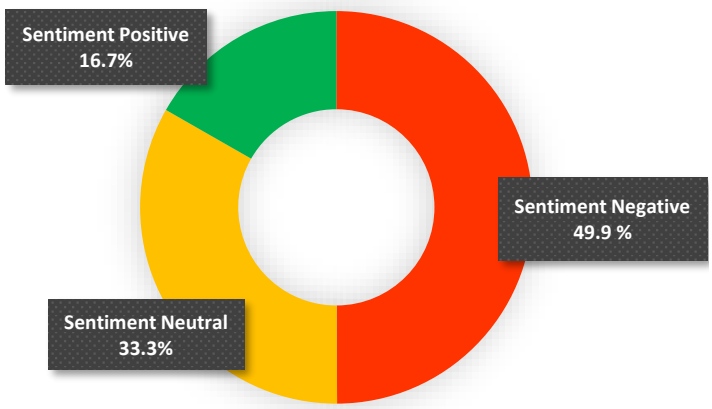
Source : Indonesian Health Profile 2010–2020

Fig.2 : Percentage of Contraceptive Use by Type in Indonesia, 2010–2020

LARCs are long-acting contraceptives that are more efficient and effective when compared to other contraceptive techniques. Long-term use is particularly effective in preventing unplanned pregnancies, making it ideal for advancing women's rights in the era of reproductive health[34]. However, the use of this product still needs to be improved to non-long-term methods. Sentiment analysis can illustrate this problem

from the public's perspective through data mining conducted on digital platforms. After mining data using web scraping techniques on Twitter social media, 4,196 tweets were obtained for the keyword "LARC" between January 1, 2020, and November 11, 2022.

Twitter provides an Application Programming Interface (API) facility as web information and data access material but with certain restrictions. Raw data from tweet mining on Twitter is filtered only to include opinions from community accounts. After that, the sentiment type classification of the incoming tweets. The classification consists of positive sentiment, negative sentiment, and neutral sentiment. The classification results will be tested with Naive Bayes and Random Forest classification methods to determine the presentation of truth and performance evaluation. The classification results of the sentiment system represent respondents' emotional tendencies on social media. Figure 3 shows the sentiment responses of Indonesians on Twitter about long-acting contraceptive products. The figure shows that more people respond with negative sentiment, which is 49.9% (2,095 tweets), while people who respond with positive sentiment are 16.7% (702 tweets). This indicates that people give more negative responses to long-term contraceptive products. The negative sentiment is suspected to be caused by myths and facts circulating in the community regarding the side effects and effectiveness of LARCs.



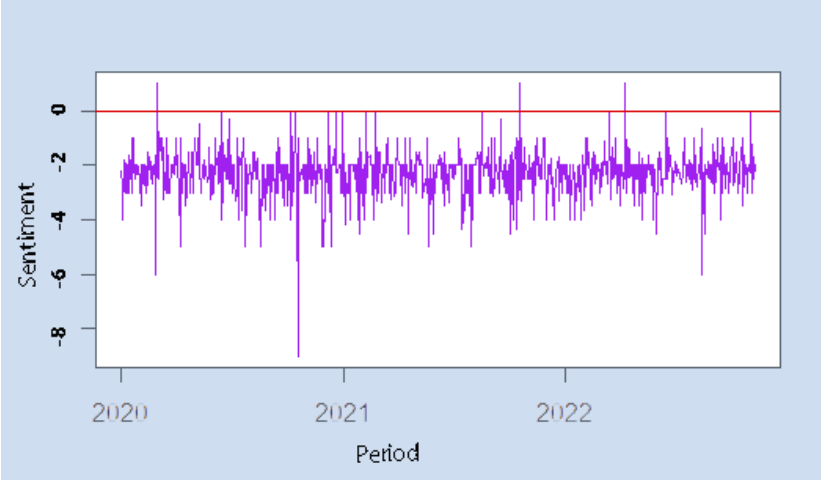
Source : Data Processing Results

Fig.2 : Percentage of Twitter Sentiment about LARC in Indonesia on January 1, 2020 - November,11 2022

Only 21.05% of women of reproductive age in Indonesia utilized LARC out of all women who used contraception. This figure is comparable to studies conducted in other countries: in Pakistan, 12.96% of individuals received postpartum IUDs; in Nepal, 6.3% of patients accepted postpartum IUDs; and in India, 9.1% of subjects had postpartum IUD insertion. It has been demonstrated that the use of LARC is still uncommon in Indonesia and other Asian countries. [35]. According to research, increased access to and use of LARC procedures is an effective tool for lowering the high incidence of unplanned pregnancy, unsafe abortion and abortion-related complications, and maternal fatalities (as well as their social and financial

impact)[9]. However, the popularity of a contraceptive method is determined by a number of circumstances. Product-related factors include availability, marketing, and media coverage of side effects and complications; provider-related factors include provider attitudes toward different methods and the prevalence of clinical personnel trained to insert and remove LARCs; and patient-related factors include marketing, access, trust, and information obtained by patients.[19]

The current issue in the context of demography is the concern that the aging population will lead to an increasing number of older people. If contraception is still applied during the aging population, it is feared that the number of children who are the next generation will decrease. Phenomena such as those that occur in developed countries such as Japan and Korea where the number of old age population is increasing while the number of children born is tiny, reducing the number of productive-age population in the country. This is a contradiction to the implementation of contraceptive programs in the current years, so the government's commitment to implementing family planning programs is slightly slackened. However, even so, there are differences between the conditions in Indonesia and developed countries such as Japan and Korea. The high rates of poverty, unemployment, school dropouts, high rates of stunting, and maternal mortality that occur in Indonesia mean that the implementation of family planning programs, especially for long-term contraceptive methods, must still be pursued so that families can do their best planning before deciding to become pregnant and have children. A more comprehensive understanding of the role of birth control policies is needed to assess current and future population aging needs and set policy priorities related to family welfare [36].



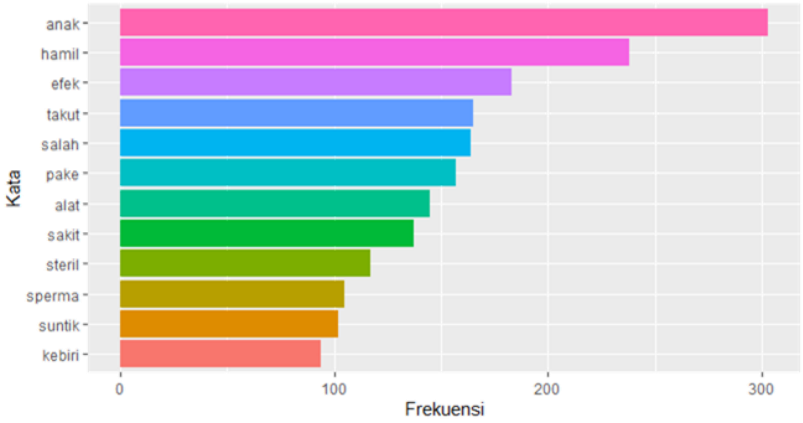
Source : Data Processing Results
Fig.3 : Sentiment on long-acting contraception from January 1, 2020 - November 11, 2022

Figure 3 shows that from January 1, 2020, to November 11, 2022, the sentiment in the Indonesian Twitter community was dominated by sentiment with a negative tone regarding long-term contraceptive methods. This negative sentiment was suspected because the period from January 1, 2020, to November 11, 2022, was the period of the COVID-19 pandemic, which caused contraceptive services to experience obstacles. The COVID-19 pandemic changed the

lifestyle of almost everyone, adversely affecting contraceptive use patterns and disrupting the distribution of medicines and contraceptives[37]. Medical workers involved in reproductive health care are diverted to meet other requirements, and consumers are hesitant to visit health facilities for reproductive health check-ups, including trips to utilize contraception. These consequences pose a substantial risk to the health and well-being of women and girls who require contraceptive services and products.The new atmosphere generated by the COVID-19 epidemic has prompted people to confront various issues. For example, the plethora of new information around COVID-19, including an overabundance of information and rumors, leading to high levels of fear; the need to learn new or so-called new normal behaviors to protect themselves from the virus; changes in typical communication patterns and trusted sources, with less personal contact and more excellent virtual communication where accessible To ensure continued access to contraception and services, including counseling and shared decisions, a number of systems are explicitly needed to maximize the use of no-touch approaches and care during the pandemic[38].

3.2. Latent Dirichlet Allocation

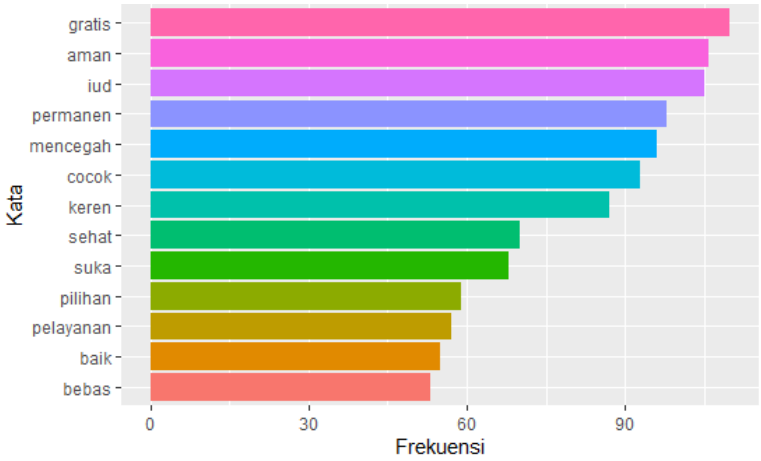
In this study, tweets discussing long-term contraceptives were grouped into three types of sentiment, as illustrated in Figure 2 and analyzed to determine which topics or words appeared most frequently. Negative sentiment can illustrate how the public responds to contraceptive products claimed to be the most effective contraception to prevent unwanted pregnancies. Through data processing using Latent Dirichlet Allocation, the grouping of words that make up the negative sentiment about long-term contraceptives can be seen in Figure 4. Based on the figure, the negative sentiment about long-term contraceptives is composed of words related to "child," "pregnant," "effect," "fear," "pain," "wrong," "sterile," and "castration" from these constituent words, it can be concluded that the negative sentiment is indicated because people's fear of using long-term contraceptives is caused by the desire to have children still and get pregnant, the effect caused by the impact of using long-term contraceptive methods, the pain caused, the mismatch or mistake in choosing the type and method of long-term contraception, and not ready to perform sterilization and castration.



Source : Data Processing Results

Fig.4 : Top Words That Make Up The Negative Sentiment Towards LARC

Public talk about long-term contraceptive methods also received a positive response. Based on Figure 5, the topics that dominate positive sentiment related to long-term contraceptive methods include the words "free," "safe," "permanent," "prevent," "suitable," "cool," "healthy," "like," "good," "service," and "free". Installing a long-term contraceptive method that is free, safe, and permanent in preventing pregnancy was analyzed as a constituent of positive sentiments about long-term contraceptive methods. In addition, positive sentiment was expressed because people liked that long-term contraceptive methods were considered suitable, gave more freedom to have sexual intercourse without fear of unwanted pregnancy, and were considered healthier. Good service in the installation of long-term contraceptive methods is also a constituent of positive sentiment in the community.



Source : Data Processing Results
Fig.5 : Top Words That Make Up The Positive Sentiment Towards LARC

3.3. Naïve Bayes and Random Forest Classification

Clustering public opinion in the form of Twitter tweets in the classification of positive, negative, and neutral sentiments can be evaluated through accuracy in machine learning models. Machine learning classification algorithms used are random forest and naive Bayes. Accuracy describes how accurately the public tweets on the Twitter channel are classified correctly based on their sentiment.

In this study, out of 4,196 tweets, 75% were used as training data and 25% as testing data. Training data as the basis for making random forest and naive bayes models. And testing data as a model evaluation to measure the level of accuracy through a confusion matrix.

Table 3. Confusion Matriks of Naive Bayes Classification

Prediction	Negative	Neutral	Positif
Negative	491	0	0
Neutral	1	330	0
Positif	31	4	181

Source : Data Processing Results

Based on Table 3, the naive Bayes model is quite balanced to predict the three sentiment models generated. Based on the table above, the model correctly predicts 181 out of 216 tweets with positive sentiment (83.7%), 330 out of 331 tweets with neutral sentiment (99.7%), and 491 out of 491 tweets with negative sentiment (100%). The model with Nave Bayes classification has an accuracy rate of 96.53%, which is classified as very good, with a confidence level of 95% in the accuracy interval between 95.23% and 97.56%. So, the Naive Bayes model can classify Twitter tweets well into three sentiment groups.

Table 4. Confusion Matriks of Random Forest Classification

Prediction	Negative	Neutral	Positif
Negative	523	0	0
Neutral	0	333	6
Positif	0	1	175

Source : Data Processing Results

Meanwhile, Table 4 shows the confusion matrix results of the random forest model. The performance of the random forest model is also quite balanced across the three models. Based on the table above, the model can correctly predict 175 out of 176 (99.7%) tweets with positive sentiment, 333 out of 339 tweets with neutral sentiment (98.2%), and 523 out of 523 tweets with negative sentiment (100%). The model has an accuracy rate of 99.3%, classified as very good, with a confidence level of 95% and an accuracy interval between 9862% and 99.7%. So, the random forest model can better classify Twitter tweets into three sentiment groups.

4. Conclusion

The sentiment analysis process is carried out to predict a phenomenon or assess products, programs, and services. In this study, the results of a sentiment analysis of public opinion regarding the use of long-term contraceptive methods are expected to provide information to relevant parties regarding how the public views contraceptive methods that are claimed to have the best level of effectiveness compared to other methods. The results of sentiment analysis taken in the period from January 1, 2020, to November 11, 2022, show that the sentiment of the Indonesian people regarding long-term contraceptive methods is still negative. Some of the things that are indicated as the cause of the negative sentiment are suspected because the data collection period was carried out during the COVID-19 pandemic when health facilities experienced obstacles in distributing contraceptives, the priority of services at health facilities is given to COVID-19 patients, and the fear felt by family planning receptors to provide services during a pandemic. In addition, through the words that build negative sentiment towards long-term contraceptive methods, it is indicated that the desire to continue to have children and become pregnant, the effects caused by the impact of using long-term contraceptive methods, the pain caused, the mismatch or error in choosing the type and method of long-term contraception, and not being ready to perform sterilization and castration are causes of negative sentiment towards long-term contraceptive methods.

The installation of long-term contraceptive methods that are free, safe, and permanent in preventing pregnancy was analyzed as a constituent of positive sentiment towards long-term contraceptive methods. In addition, positive sentiment was expressed because people liked that long-term contraceptive methods were considered suitable, provided freedom in sexual intercourse without fear of unwanted pregnancy, and were considered healthier. Good service in installing long-term contraceptive methods is also a constituent of positive sentiment in the community. The sentiment

classification results from Twitter tweets are considered very good because both Nave Bayes and random forest methods provide accuracy above 90%, namely 96.53% for the Naive Bayes method and 99.3% for the random forest method.

The results of this study provide advice to the government and related agencies to continue to make socialization efforts with the public regarding the effectiveness of long-term contraceptive methods in preventing unwanted pregnancies. In addition, health facility services need to be improved so that consumer satisfaction with long-term contraceptive method users can continue to increase. This study still has limitations, among others, regarding the need for further research to see the factors that cause negative sentiment in the community regarding long-term contraceptive methods, both factors from the individual side, which include social, demographic, and economic factors, as well as contextual factors that come from education, information, and the quality and quantity of family planning services, especially long-term contraceptive programs.

Reference

1. Statistic of Indonesia: Hasil Long Form Sensus Penduduk 2020. (2020)
2. Fauzia, R., Kurniawan Sujono, F.: Crisis And Communication Management In National Population And Family Planning Board In The Post-Reform Era. https://ejournal.ipdn.ac.id. 7, 1–12 (2017)
3. Taebenu, M.M.: Stalled Fertility Decline In East Nusa Tenggara, Indonesia, 2002-2017: Understanding Its Determinants. Jurnal Kependudukan Indonesia. 15, 227 (2021). https://doi.org/10.14203/jki.v15i2.504
4. United States Agency for International Development Bureau for Global Health Office of Population and Reproductive Health: WHO updates recommendations to guide family planning decisions. (2022)
5. Pazol Karen, Ellington, S.R., Fulton, A.C., Zapata, L.B., Boulet, S.L., Rice, M.E., C., Romero, L., Lathrop, E., Hurst, S., Kroelinger, C.D.: Contraceptive Use Among Women at Risk for Unintended Pregnancy in the Context of Public Health Emergencies - United States, 2016. MMWR Morb Mortal Wkly Rep. 67, 898–902 (2016)
6. Lawrence B Finer, Mia R Zolna: Declines in Unintended Pregnancy in the United States, 2008-2011. N Engl J Med. 374, 843–852 (2016)
7. Gariepy, A.M., Lundsberg, L.S., Miller, D., Stanwood, N.L., Yonkers, K.A.: Are pregnancy planning and pregnancy timing associated with maternal psychiatric illness, psychological distress and support during pregnancy? J Affect Disord. 205, 87–94 (2016). https://doi.org/https://doi.org/10.1016/j.jad.2016.06.058
8. Weisberg, E.: Citation: Weisberg E. Promoting the Use of Long-Acting Reversible Contraceptives. (2014)
9. Bahamondes, L., Fernandes, A., Monteiro, I., Bahamondes, M.V.: Long-acting reversible contraceptive (LARCs) methods. Best Pract

Res Clin Obstet Gynaecol. 66, 28–40 (2020). https://doi.org/https://doi.org/10.1016/j.bpobgyn.2019.12.002
10. Yuliati, I.: Peramalan Dan Analisis Hubungan Faktor Penggerakan Lini Lapangan Dalam Meningkatkan Peserta KB Aktif MKJP. (2021)
11. Newman, N., Fletcher, R., Schulz, A., Andı, S., Robertson, C.T., Kleis Nielsen, R.: The Reuters Institute Digital News Report 2021. (2021)
12. Brian Dean: How Many People Use Twitter in 2023? [New Twitter Stats], https://backlinko.com/twitter-users, (2023)
13. Sitaram Asur, Bernardo A., Huberman: Predicting the Future with Social Media. Appl Energy. 112, 1536–1543 (2013). https://doi.org/10.1016/j.apenergy.2013.03.027
14. Carter, M.W., Bergdall, A.R., Henry-Moss, D., Hatfield-Timajchy, K., Hock-Long, L.: A qualitative study of contraceptive understanding among young adults. Contraception. 86, 543–550 (2012). https://doi.org/10.1016/j.contraception.2012.02.017
15. Ali, M.M., Amialchuk, A., Dwyer, D.S.: Social Network Effects in Contraceptive Behavior Among Adolescents. Journal of Developmental & Behavioral Pediatrics. 32, 563–571 (2011). https://doi.org/10.1097/DBP.0b013e318231cf03
16. Matt G. Southern: Twitter Releases Study Showing Exposure To Brand Tweets Drives Action Online And Offline, https://www.searchenginejournal.com/twitter-releases-study-showing-exposure-brand-tweets-drives-action-online-offline/104099/#close, (2014)
17. Merz, A.A., Gutiérrez-Sacristán, A., Bartz, D., Williams, N.E., Ojo, A., Schaefer, K.M., Huang, M., Li, C.Y., Sandoval, R.S., Ye, S., Cathcart, A.M., Starosta, A., Avillach, P.: Population attitudes toward contraceptive methods over time on a social media platform. In: American Journal of Obstetrics and Gynecology. pp. 597.e1-597.e14. Mosby Inc. (2021)
18. Kofinas, J.D., Varrey, A., Sapra, K.J., Kanj, R. V., Chervenak, F.A., Asfaw, T.: Adjunctive Social Media for More Effective Contraceptive Counseling. Obstetrics & Gynecology. 123, 763–770 (2014). https://doi.org/10.1097/AOG.0000000000000172
19. Merz, A.A., Gutiérrez-Sacristán, A., Bartz, D., Williams, N.E., Ojo, A., Schaefer, K.M., Huang, M., Li, C.Y., Sandoval, R.S., Ye, S., Cathcart, A.M., Starosta, A., Avillach, P.: Population attitudes toward contraceptive methods over time on a social media platform. Am J Obstet Gynecol. 224, 597.e1-597.e14 (2021). https://doi.org/10.1016/j.ajog.2020.11.042
20. Haerul Jaman, J., Abdulrohman, R.: Sentiment Analysis of Customers on Utilizing Online Motorcycle Taxi Service at Twitter with the Support Vector Machine. (2019)
21. Mohamed Zakir H. and Vinila Jinny, S.: A Comparative Study on Data Cleaning Approaches in Sentiment Analysis. In: Jayakumari J. and

Karagiannidis, G.K. and M.M. and H.S.A. (ed.) Advances in Communication Systems and Networks . pp. 421–431. Springer Singapore, Singapore (2020)
22. Javed, M., Kamal, S.: Normalization of Unstructured and Informal Text in Sentiment Analysis. (2018)
23. Ela Kumar: Natural Language Processing. I.K International Publishing House Pvt.Ltd, New Delhi (2010)
24. Al-Saqqa, S., Awajan, A., Ghoul, S.: Stemming Effects on Sentiment Analysis using Large Arabic Multi-Domain Resources. Presented at the August (2019)
25. S Vijayarani, R Janani: Text Mining: open Source Tokenization Tools – An Analysis. Advanced Computational Intelligence: An International Journal (ACII). 3, 37–47 (2016). https://doi.org/10.5121/acii.2016.3104
26. Avinash M. and Sivasankar, E.: A Study of Feature Extraction Techniques for Sentiment Analysis. In: Abraham Ajith and Dutta, P. and M.J.K. and B.A. and D.S. (ed.) Emerging Technologies in Data Mining and Information Security. pp. 475–486. Springer Singapore, Singapore (2019)
27. Pramita, D., Saptono, R., Anggrainingsih, R.: Academic Articles Classification Using Naive Bayes Classifier (NBC) Method. ITSMART: Jurnal Teknologi dan Informasi. 7, (2018). https://doi.org/https://doi.org/10.20961/itsmart.v7i2.23844
28. Nusiyono, A.J., Huda, Q.: Sentiment Analysis Of Personal Data Protection With Machine Learning Approach. Jurnal Pertahanan & Bela Negara. 13, (2023)
29. Cutler, A., Cutler, D.R., Stevens, J.R.: Random Forests. In: Ensemble Machine Learning. pp. 157–175. Springer New York, New York, NY (2012)
30. Breiman, L.: Random Forests. Mach Learn. 45, 5–32 (2001). https://doi.org/10.1023/A:1010933404324
31. Sheth, V., Tripathi, U., Sharma, A.: A Comparative Analysis of Machine Learning Algorithms for Classification Purpose. Procedia Comput Sci. 215, 422–431 (2022). https://doi.org/https://doi.org/10.1016/j.procs.2022.12.044
32. Fajarningtiyas, D.N., Sulistiawan, D., Naibaho, M.M.P., Arifa, R.F.: Pattern and determinant of contraceptive use among women in indonesia from 2007 to 2017: Evidence from demographic and health survey. Open Access Maced J Med Sci. 9, 1363–1370 (2021). https://doi.org/10.3889/oamjms.2021.7417
33. Indonesian Ministry of Health: Indonesia Health Profile. , Jakarta (2020)
34. Kungu, W., Khasakhala, A., Agwanda, A.: Use of long-acting reversible contraception among adolescents and young women in

Kenya. PLoS One. 15, e0241506 (2020). https://doi.org/10.1371/journal.pone.0241506
35. Andardi, B., Rahim, D.G.S.P., Achadi, A.: Reasons of Refusal to Long Acting Reversible Contraception (LARC) on Reproductive Age Women: a Scoping Review. e-CliniC. 10, 364 (2022). https://doi.org/10.35790/ecl.v10i2.41489
36. Shu Binbin and Li, J.X.: Birth Control Policy and Population Aging. In: Gu Danan and Dupre, M.E. (ed.) Encyclopedia of Gerontology and Population Aging. pp. 1–5. Springer International Publishing, Cham (2019)
37. Chris Purdy: Opinion: How will COVID-19 affect global access to contraceptives — and what can we do about it?, (2020)
38. Stephanie Desmon: Family Planning Still Needed During Pandemic, (2020)