International Journal of Integrated Science and Technology (IJIST)
Vol.3, No.3, 2025: 1391-1398

Analysis of Public Sentiment on Child Development Impact of Divorce With Naïve Bayes Algorithm Data Classifier on Twitter Using Rapidminer

Nafar Dwi Wijaya[1], Suherman[2], Asep Suprianto[3]
Universitas Pelita Bangsa
Corresponding Author: Nafar Dwi Wijaya: nafardwiwijaya18@gmail.com

A R T I C L E I N F O

Keywords: Sentiment Analysis, Divorce, Child Development, Naive Bayes Classification, RapidMiner, and Twitter.

Received : 12, January
Revised : 20, February
Accepted: 25, March

A B S T R A C T

Divorce is a social phenomenon that can affect children's psychological, emotional and social development. Social media such as Twitter has developed into an important platform for people to share their opinions and experiences on these issues in today's internet era. As a result, an analysis technique is needed that can automatically process and classify people's feelings. This study uses the Naive Bayes Classifier algorithm to analyze public sentiment regarding child development caused by divorce. The data used is taken from Twitter via scraping, which then goes through preprocessing processes such as cleaning, tokenization and stemming. This makes the data cleaner and ready for analysis. Next, the data is categorized based on sentiment: positive, negative, and neutral. RapidMiner software is used to carry out the analysis process, which makes it easier to use the Naïve Bayes algorithm in sentiment classification. The research results show that the Naïve Bayes algorithm has a fairly high level of accuracy in classifying public sentiment regarding the impact of divorce on children. The results of this analysis show that public opinion tends to be more negative, indicating concern about the condition of children after divorce. This research hopes to help governments, psychologists and social institutions create better policies and programs to help children affected by divorce.

INTRODUCTION

Parents act as primary educators and families are the first environment that shapes a child's personality. A harmonious family is very important to build children's character (Ajrina, 2021). However, not all families are able to maintain harmonious relationships, so divorce becomes a big problem for children (Yani & Zakiyah, 2005). Children who experience divorce can experience difficulties in social and emotional development, loss of parental affection, and mental distress. Some children adapt well, but the unstable family environment causes psychological disorders in many people (Ismiati, 2018). Negative behavior in the surrounding environment, including at school, is often associated with this state. In today's internet age, Twitter has evolved into a place to discuss the impact of divorce on children (Wardani et al., 2022). Sentiment analysis, which is the process of processing and assessing individual emotions or attitudes towards a problem, can be used to analyze opinions spread on social media (Kenny Syalwatyarsa & Mirna Nur Alia Abdullah, 2024). In this study, the Naive Bayes approach was used to classify public sentiment regarding post-divorce child development (Amalia Yunia Rahmawati, 2020). This algorithm can not only identify positive, negative, and neutral opinions, but can also provide an overview of public perception (Diara Eka Yogianti, 2019).

THEORETICAL FRAMEWORK

A. Sentiment Analysis

Sentiment analysis is widely used in marketing, customer service, and politics to understand public opinion. It is the process of identifying opinions and emotions in texts such as reviews, comments, and social media posts to determine whether they are positive, negative, or neutral. Sentiment analysis can automatically measure customer satisfaction, assess brand reputation, and find public opinion trends from various sources, such as e-commerce and social media, thanks to the use of machine learning and natural language processing (NLP) techniques.

B. *The Naïve Bayes Classifier*

The Naive Bayes Classifier (NBC) is a classification algorithm based on Bayes' theorem with independent assumptions, which is used to predict data as accurately as possible. In applying the Naive Bayes Classifier method, the following equations are used:

$$P(Wi/Cj) = \frac{N_{cw} +}{N_C + V}$$

Where:

- Ncw is the number of occurrences of the word Wi in Class Cj.
- Nc is the total number of words in the class Cj.
- V is the measure of the entire vocabulary

C. *Data Mining*

Data mining is a series of actions to determine relationship patterns and extract additional value from both important informational data, such as knowledge. The purpose of data mining is to determine relationships and simplify data so that information can be understood and used properly. This process is aided by mathematics, statistics, artificial intelligence, and machine learning.

D. *Rapidminer*

The company of the same name created the data science software platform RapidMiner, which offers a uniform environment for machine learning (for predictive analytics, deep learning, text mining, machine learning and analytics) [10]. This program, known as RapidInner, is open source and available to all users. RapidMiner is used as a solution to evaluate data handling. Many approaches, including descriptive and predictive approaches, are used in RapidMiner.

METHODS

The purpose of this study was to create and apply the Naive Bayes Classifier to assess public opinion about the impact of divorce on children based on sentiment that emerged on Twitter. This model determines whether the sentiment is positive or negative. Using machine learning-based sentiment analysis, the study emphasizes the importance of understanding public opinion to support more responsive public health policies and increase awareness of child victims of divorce.

1. Data Collection

The data used in this study is Twitter text data collected using the Twitter API. The data collected included tweets containing keywords related to divorce and child growth and development within a certain period of time. The stages of data collection include:

1. Keyword determination: 4 keywords are used, including "divorce", "divorced child", "due to divorce", and other variations.
2. Data scraping: Collect tweet data in CSV format using Twitter API.
3. Pre-Data Processing: Remove duplicates, irrelevant tweets, and tweets with inappropriate language.

4. Pre - Processing Data

This stage converts raw data into data that can be used for sentiment analysis. Initial Pre-Processing includes:

1. Cleansing Data

Eliminates unnecessary punctuation, numbers, and special characters.

2. Tokenization

Divide text into individual words.

3. Stopword Removal

Omit common words that don't provide important information (e.g. "in," "to," "and").

4. Vote

Changing words to their basic form (e.g. "ran" to "ran").

5. Data Normalization

Changing non-standard words to standard words.

5. Labelling Data

After the data is cleaned, a data labeling process is carried out to distinguish positive, negative, and neutral sentiments. This labeling can be done manually or you can use a sentiment dictionary to automate labeling.

6. Data Sharing

Labeled data is divided into two sets:

1. Training Data: 80% of the total data is used to train the model.
2. Test Data: 20% of the total data is used to test the accuracy of the model created.

1. Data Types

The data used in this study is unstructured text data from Twitter users' tweets about divorce that has an impact on child development used as a source of research data. A number of criteria, including neutral, negative, and positive emotions are used to evaluate each tweet. The Naive Bayes Classifier method is used in the analysis to estimate the way individuals express their opinions. RapidMiner's technology is used to process this data to accurately classify opinions expressed on Twitter about child development as a result of divorce.

2. Data Used

This research dataset, which includes public comments on child development impacted by divorce, was obtained via Twitter. This dataset consists of tweets taken from the Twitter network and classified as neutral, positive or negative. After preprocessing the tweet to exclude unnecessary components such as emoticons, URLs, and punctuation, the Naive Bayes Classifier algorithm from the RapidMiner platform is used to evaluate the data.

3. Model Testing

To ensure the effectiveness and dependency of the model in categorizing positive, negative, or neutral sentiment, the Naïve Bayes Classifer model went through several stages of testing while examining public opinion regarding child development as a result of divorce on Twitter. The first step starts with 80:20 data between training and testing data. The test data is used to assess the performance

of the trained model, while the training data is used to train the Naïve Bayes model.

Metrics including accuracy, precision, recal, and F1-score are used to assess the model's performance. These measures provide a broad picture of how successfully algorithms use Twitter data to categorize public opinion about child development and the impact of divorce.

RESULTS

Naïve Bayes

1. Gathering information from data

- Total number of test data: 181
- True label distribution:
  - Negative = 49
  - Positive = 61
  - Neutral = 31
- Model Prediction:
  - Negative Prediction: 49 true, 0 false.
  - Positive Prediction: 59 true, 2 false.
  - Neutral Prediction: 27 true, 2 false.
- Model accuracy: 95.74%
- Recall per class:
  - Negative: 92.45%
  - Positive: 96.72%
  - Neutral: 100%

2. Calculating Prior Probability

The formula of prior probability in Naïve Bayes is:

$$P(C) = \frac{Number\ of\ data\ class\ C}{Total\ data}$$

So we calculate:

- **P(Negative):**

$$P(\text{Negative}) = \frac{49}{181} = 0.2727\ (27.07\%)$$

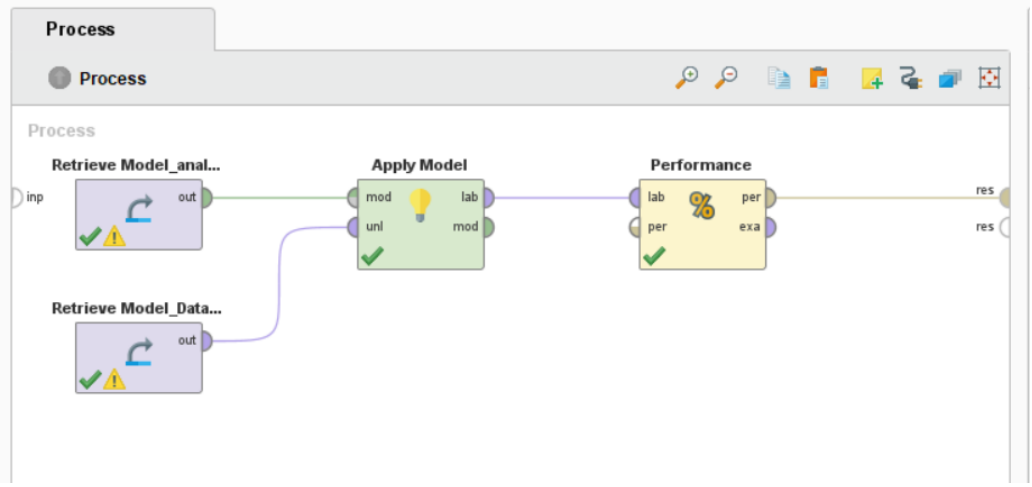
- **P(Positive):**

$$P(\text{Positive}) = \frac{61}{181} = 0.3365\ (33.65\%)$$

- **P(Neutral):**

$$P(\text{Neutral}) = \frac{31}{181} = 0.1713\ (17.13\%)$$

Table 1. 1 Rapidminer's Naïve Bayes Classification Testing Model



As shown above, Naive Bayes classification modeling classifies them into two categories: positive negative and neutral. The algorithm in this study produced an accuracy of 95.74% using the Naive Bayes method from testing data that had been collected and tested with training and data testing.

Testing the results of the data with a rapid miner will result in a precession class and a recall class. Precesion is the comparison of the number of relevant documents with the total documents of the result of the question, and recall is the comparison of the number of relevant documents with the total relevant documents. In addition to accuracy, there are two things that are generated in the classification of texts, so that the testing of the data can be seen in the image.

- Accuracy: Naive Bayes models reach 95.74%.
- Accuracy and Recall

Table 1. 2 Rapidminer Harvesting Results

Table View Plot View

accuracy: 95.74%

	true Negatif	true Posistif	true Neutral	class precision
pred. Negatif	49	0	0	100.00%
pred. Posistif	2	59	0	96.72%
pred. Neutral	2	2	27	87.10%
class recall	92.45%	96.72%	100.00%	

Class Precision: Negative: 100.00%, Positive: 96.72%, Neutral: 87.10%

Class Recall: Negative: 92.45%. Positive: 96.72%, Neutral: 100.00%

Tweet data test results with Rapid miner with the classification of the naïve Bayes method.

DISCUSSION

The results of the analysis of public sentiment about the impact of divorce on child development. The Naive Bayes model manages to place sentiment in different categories, with negative sentiment being the most common. The importance of understanding public perceptions when planning solutions to the social problems caused by divorce is underlined in the discussion.

CONCLUSIONS AND RECOMMENDATIONS

The results of the analysis show that public opinion on Twitter is dominated by negative opinions about the impact of divorce on children, indicating concerns about emotional and social disorders in children.

The Naive Bayes algorithm has proven to be effective in analyzing sentiment. The results achieved 95.74% accuracy, 100.00% negative precision, 96.72% positive, 87.10% neutral, and 92.45% negative recall, 96.72% positive, and 87.10% neutral. This algorithm works well for both positive and negative categories.

These findings help stakeholders make decisions about public perceptions and make policies that are more in line with needs.

FURTHER STUDY

Expansion of Data Sources: Further research can use data from other social media platforms, such as Facebook and Instagram, to broaden perspectives. Improved Preprocessing Techniques: Using certain techniques to deal with ambiguous contexts and improve the accuracy of the analysis. Use of Additional Algorithms: Check the performance of algorithms such as SVM or deep learning. Temporal Analysis: Sentiment analysis is used to find out how people's opinions change over a period of time. Demographic Aspects: Variables such as age, gender, and location are included for further analysis. Collaboration from different disciplines: Engaging experts from different fields to produce more comprehensive research.

ACKNOWLEDGMENT

The author would like to thank everyone who has helped and supported me in completing this thesis. Thank you also to all lecturers and colleagues at Universitas Pelita Bangsa who have provided valuable advice and criticism to improve the quality of this work. Additionally, I would like to thank my family members for their support and understanding during the writing process. The completion of this project would not have been possible without the contribution and support of all these parties.

REFERENCES

Ajrina, A. (2021). The impact of parental divorce on children's social behavior in West Pontianak District, West Kalimantan. Journal of S-1 Sociology, 3(3), 1–19. https://fisipuntan.org/jurnal/index.php/sociologique/article/view/1232

Amalia Yunia Rahmawati. (2020). The Impact of Broken Home on Social-Emotional Development. July, 1–23.

Diara Eka Yogianti. (2019). Parental Divorce and Its Impact on Adolescent Emotional Development in Hargomulyo Village, Sekampung District. Thesis. https://repository.metrouniv.ac.id/id/eprint/3574/1/SKRIPSI BY EKA YOGIYANTI NPM. 14114011.pdf

Ismiati, I. (2018). Parental Divorce and Child Psychological Problems. At-Taujih: Islamic Guidance and Counseling, 1(1), 1–16. https://doi.org/10.22373/taujih.v1i1.7188

Kenny Syalwatyarsa, & Mirna Nur Alia Abdullah. (2024). Divorce in the Digital Era: Analyzing the Factors and Impact of Cheating Trends in the Perspective of Social Demographics. SABANA: Journal of Sociology, Anthropology, and Nusantara Culture, 3(1), 39–48. https://doi.org/10.55123/sabana.v3i1.3310

Wardani, A. K., Suhariadi, F., & Sugiarti, R. (2022). The Impact of Divorce on Children's Social Behavior. Citizenship, 6(2), 2684–2690.

Yani, & Zakiyah, T. (2005). Background and Impact of Divorce (Case Study at the Wonosobo Religious Court). 50–51.