International Journal of Advances in Data and Information Systems
Vol. 2, No. 1, April 2021, pp. 36~44
ISSN: 2721-3056, DOI: 10.25008/ijadis.v2i1.1216

Sentiment Analysis Approach for Analyzing iPhone Release using Support Vector Machine

Wasim Bourequat[1], Hassan Mourad[2]
[1,2]Department of Computer Engineering, Université Internationale de Casablanca

Article Info

Article history:

Received Feb 12, 2021
Revised Mar 19 2021
Accepted Apr 03, 2021

Keywords:

Sentiment Analysis
Support Vector Machine
SVM
Machine Learning
Classification

ABSTRACT

Sentiment analysis is a process of understanding, extracting and processing textual data automatically to get sentiment information contained in a comment sentence on Twitter. Sentiment analysis needs to be done because the use of social media in society is increasing so that it affects the development of public opinion. Therefore, it can be used to analyze public opinion by applying data science, one of which is Natural Language Processing (NLP) and Text Mining or also known as text analytics. The stages of the overall method used in this study are to do text mining on the Twitter site regarding iPhone Release with methods of scraping, labeling, preprocessing (case folding, tokenization, filtering), TF-IDF, and classification of sentiments using the Support Vector Machine. The Support Vector Machine is widely used as a baseline in text-related tasks with satisfactory results, on several evaluation matrices such as accuracy, precision, recall, and F1 score yielding 89.21%, 92.43%, 95.53%, and 93.95, respectively.

Corresponding Author:

Wasim Bourequat,
Department of Computer Engineering,
Université Internationale de Casablanca
Casablanca 50169, Maroko
Email: wbourequat@gmail.com

1. INTRODUCTION

The rapid development of various web portals and social media has made the availability of textual document sources very large and easily accessible. One example of social media that is widely used is Twitter. Through this social media, users write various kinds of opinions, experiences and other matters of concern every day. This article is often referred to as a tweet. With the development of technology, tweets sent by Twitter users can already analyze the polarity of their sentiments.

The growth of information technology, especially social media, has changed the way humans communicate with each other. The use of social media is widely used by the general public. The public uses social media a lot to express their opinions, experiences and other things that concern them [1]. These kinds of things are often called sentiments.

Nowadays, the increasing use of social media in society will influence the development of public opinion along with the increasing role of sharing information on social media. Therefore, it can be used to analyze public opinion whether it is positive, negative or neutral. The technique for analyzing public opinion is called sentiment analysis. Sentiment analysis is a field of study that analyzes a person's opinion, opinion, evaluation, judgment, attitudes and emotions towards entities such as products, services, organizations, individuals, problems, events, topics and attributes [2]. Sentiment analysis is a type of natural language processing, which is word processing to track people's moods about certain products or opinions. Sentiment analysis is a technique for detecting

opinions on a subject (for example, individuals, organizations or products) in a data set [3]. Also, opinion is the expression of an attitude about an issue that contains an attitude about issues that contain contradictions [4]. Opinions here can be in the form of mentions on social media or articles on news sites and personal blogs.

Sentiment analysis has been used extensively on various problems in the real world and is implemented using various algorithms [1], [5], [6]. The data can be sourced from various social media platforms, such as twitter, facebook, instagram, youtube, tripadvisor, and so on [7]–[10].

With this phenomenon of abundant data, many people every day deal with data that comes from various types of observations and measurements. One method of utilizing this abundant data is Text Mining. Text mining is a method used to extract quality information from a series of texts summarized in a document. Text Mining can be used to find out the required information such as public sentiment on something. From these sentiments, we can find out the responses from the crowd such as positive or negative sentiments. With the current popularity of Twitter, it is certainly an ideal place for researchers to dig up information about the public's response, one of which is related to the iPhone release on Twitter.

Regarding the release of the iPhone, what is considered by the public to judge a new product from Apple is getting a good response from the public or not, because usually people also give their opinion about the product. However, data volume problems, such as the number of comments that can reach hundreds or even thousands of comments, cause manual comment analysis to take time and effort. In addition, comments that do not follow standard sentences are also a challenge. Therefore, this is what motivates this research to build a sentiment analysis system for Twitter user comments about the iPhone release automatically, quickly and accurately.

2. LITERATURE REVIEW

2.1 Twitter

Twitter is a social media and microblogging service that allows users to send realtime messages. This message is popularly known as a tweet. A Tweet is a short message with a length of characters limited to 140 characters. Due to the limited character that can be written, a tweet often contains abbreviations, slang language or spelling errors [6]. Since its inception, Twitter was created as a mobile-based service designed according to the character limits of a text message (SMS), and to this day, Twitter can still be used on any cell phone that has the ability to send and receive text messages.

Twitter was created to be a place to share experiences between fellow users without any barriers. By using it, users will find it easy to follow trends, stories, information and news from all corners of the world. Apart from that, Twitter also helps its users to stay connected with the people closest to them. When a user sends a tweet, the message is public and can be accessed by anyone, anywhere and anytime. In fact, for people who follow this Twitter account, the tweet will automatically appear in that person's timeline.

The following are some of the terms that are known on Twitter:

1. Mention. Mention is mentioning or calling other Twitter users in a tweet. Mention is done by typing '@' followed by another username.
2. Hashtags. Hashtags are used to tag a topic of conversation on Twitter. Hashtag writing starts with a '#' followed by the topic being discussed. Hashtags are commonly used to increase the visibility of users' tweets.
3. Emoticons. Emoticons are facial expressions that are represented by a combination of letters, punctuation and numbers. Ordinary users use emoticons to express the mood they are feeling.
4. Trending topics. If hashtags are a way to tag a topic of conversation on Twitter, then trending topics are a collection of topics that are very popular on Twitter.

A surprising fact was presented by Giummole et al. [11] stated that trending topics on Twitter will increase the prediction of search results on Google. This can happen because in every five minutes, Twitter will issue a list of very popular topics (trending topics) by monitoring and analyzing conversations, while Google will also issue a list of popular searches that are searched by its users every hour.

The development of Twitter to date has offered the opportunity to study human cultural behavior like never before. This can be done because Twitter has prepared defined data (for example XML and category lists) as well as unstructured data (for example in the tweet text) [2], [12]. Twitter has a structured model [2] as well as unstructured text [2], [12]. Structured text can be seen in the implementation of metadata in each tweet, where there is always information on username, timestamp, text, retweet, favorites and other information. Unstructured text can be seen in the text section or more commonly referred to as tweets because the content in this section does not have a special structure. The only rule that applies is the maximum number of characters that can be entered is 140 characters [13], [14].

2.2 Sentiment Analysis

According to Liu [2], sentiment analysis is a field of study that analyzes people's opinions, opinions, evaluations, judgments, attitudes and emotions towards entities such as products, services, organizations, individuals, problems, events, topics and attributes. Meanwhile, according to Haddi, Liu and Shi [15], sentiment analysis in reviews is the process of investigating product reviews on the internet to determine opinions or feelings about a product as a whole.

Sentiment analysis is a very interesting field to be developed in the digital world because at this time the community generally expresses their feelings of opinion and also the results of their thoughts through cyberspace with text language where readers sometimes have misunderstandings in translating the sentiments contained therein. By analyzing public sentiment, of course, it will produce new information that can be processed to produce useful information as well.

Opinion texts are included in unstructured data, so preprocessing is necessary to make the data structured and can be processed to take existing aspects through tokenization, word segmentation, Part-of-Speech Tagging, stemming, and others.

In general, sentiment analysis is divided into three levels, namely the document level, the sentence level, and the fine-grained level. Document-level and sentence-level can also be categorized into coarse-grained levels. Methods in sentiment analysis are divided into two types, namely learning-based and lexical-based. Learning-based uses training data and testing data, while lexical-based uses a dictionary (opinion lexicon).

2.3 Text Mining

Text Mining is the process of extracting patterns (information and knowledge) from a number of unstructured data. There are two inputs for text mining, namely unstructured data (Word documents, PDFs, text citations) and structured data [16].

Text Mining is used to answer business questions and to optimize the efficiency of day-to-day operations and improve long-term strategic decisions. Text mining is often used in techniques such as categorization, entity extraction and sentiment analysis which are used to identify insights and trend patterns in large volumes of structured data.

Social media, which is a potential source of structured data, is considered a valuable source of market and customer intelligence information. Many companies use text mining to analyze or predict customer needs and assess their brand perception. Text analysis can address these issues by analyzing large volumes of structured data, expressing opinions, emotions and sentiments and their relationship to brands and products.

2.4 Machine Learning

Machine learning is one of the fields of science in Artificial Intelligence. Machine learning aims to make machines trained with many examples or datasets related to the required task. The machine learns the given patterns based on the dataset and generates a rule of its own. So that when data is entered into the machine, the machine can recognize the data. In general, machine learning is divided into four broad categories, namely supervised learning, unsupervised learning, self-supervised learning, and reinforcement learning [17], [18].

Supervised learning is the approach used most often. Supervised learning creates machine learning from labeled or annotated datasets. Whereas unsupervised learning is the opposite, by

providing a dataset that is not labeled. Selfsupervised learning is a supervised learning but without a dataset labeled by annotators. The dataset used still uses labels but labels are obtained from data input using a heuristic algorithm [17].

2.5 TF-IDF

The TF-IDF method is a method for calculating the weight of each word that is most commonly used in information retrieval. This method is also known to be efficient, easy and has accurate results [19], [20]. The Term Frequency-Inverse Document Frequency (TF-IDF) method is a way of giving weight to the relationship of a word (term) to a document. TF-IDF is a statistical measure used to evaluate how important a word is in a document or in a group of words. For a single document each sentence is considered a document. The frequency with which the word appears in a given document shows how important it is in the document. The number of times a document contains this word indicates how common it is. The weight of the word is greater if it appears in a document frequently and it is smaller if it appears in multiple documents [19], [20].

In the TF-IDF algorithm, a formula is used to calculate the weight (W) of each document against keywords with the formula, namely:

$$W_{t,d} = tf_{t,d} \times idf_t = tf_{t,d} \times \log \frac{N}{df_t} \quad (1)$$

where

$W_{t,d}$ = weights of TF-IDF
$tf_{t,d}$ = number of word frequencies
idf_t = inverse number of document frequency per word
df_t = number of document frequency per word
N = total number of documents

2.6 Support Vector Machine

Support Vector Machine (SVM) is a set of guided learning methods (supervised learning) that analyzes data and recognizes patterns, used for classification and regression analysis (Saraswati, 2011).

Support Vector Machine is a machine learning method that works on the principle of structural risk minization (SRM) with the aim of finding the best hyperlane that separates two classes in the input space. $x^k = f^k, and\ w, b$ are obtained by minimizing the loss function.

$$L(w,b) = w^T w + c \sum max\left(0{,}1 - y^i\left(w^T F^{(i)} + b\right)\right)^2 \quad (2)$$

This classification is done by looking for a hyperplane or decision boundary that separates one class from another [21]. In this concept, Support Vector Machine tries to find the best hyperplane among the unlimited number of functions. The unlimited function in hyperplane search in the Support Vector Machine method is an advantage, where processing can always be done no matter what data it has. The hyperplane can be seen in Figure 1.

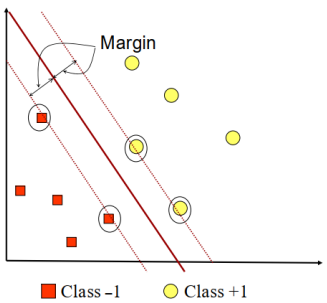

Figure 1. Hyperlane of SVM

2.7 Matrix Evaluation

The test is carried out by measuring the accuracy of the method and the suitability of the system being built. In testing the performance of the classification method, it is measured using an evaluation matrix which includes accuracy, precision, recall and f1-score. The table regarding the actual and predictive class evaluation matrix [22] can be seen in Table 1 below.

Confusion Matrix		**Predicted Class**	
		Class = Yes	**Class = No**
Actual Class	**Class = Yes**	TP	FN
	Class = No	FP	TN

The confusion matrix consists of four categories, namely True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). True Positive (TP) is a sentence that has a positive sentiment and the prediction results also show a positive sentiment. False Positive (FP) is a sentence that has positive sentiment but the prediction results show negative sentiment. True Negative (TN) is a sentence that has negative sentiments and the prediction results also show negative results. False Negative (FN) is a sentence that has a negative sentiment but the prediction results show a positive sentiment.

After obtaining values for confusion matrix, accuracy, precision, recall, and F-measure values can also be obtained. Accuracy aims to show the percentage of input that SVM successfully predicts. Precision aims to calculate the percentage of the input detected by the system, for example, the system labels the input as positive which is originally positive as well. Recall aims to calculate the percentage of input that the system identifies as true as true. Meanwhile, the F-measure is the average obtained from precision and recall. The calculation formula to get accuracy is shown in equation 3. Meanwhile, the formula for calculating precision and recall is carried out for each sentiment with examples such as in equations 4 and 5.

$$Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (3)$$

$$Precision = (TP)/(TP + FP) \quad (4)$$

$$Recall = (TP)/(TP + FN) \quad (5)$$

2.8 ROC Curve

ROC curves show accuracy and visually compare classifications. The ROC curve expresses confusion matrix. The ROC is a two-dimensional graph with false positive as a horizontal line and true positive as a vertical line. General guidelines for classifying test accuracy using the AUC [23], [24].

ROC curves show accuracy and compare classifications visually. The ROC curve expresses confusion matrix. The ROC is a two-dimensional graph with false positive as a horizontal line and true positive as a vertical line. General guidelines for classifying test accuracy using the AUC.

0.90 - 1.00 = Excellent Classification;
0.80 - 0.90 = Good Classification;
0.70 - 0.80 = Fair Classification;
0.60 - 0.70 = Poor Classification;
0.50 - 0.60= Failure.

3. RESEARCH METHOD

Research methodology is a general description of how the system to be built in this study works. The research methodology in this study can be seen in Figure 2 below.

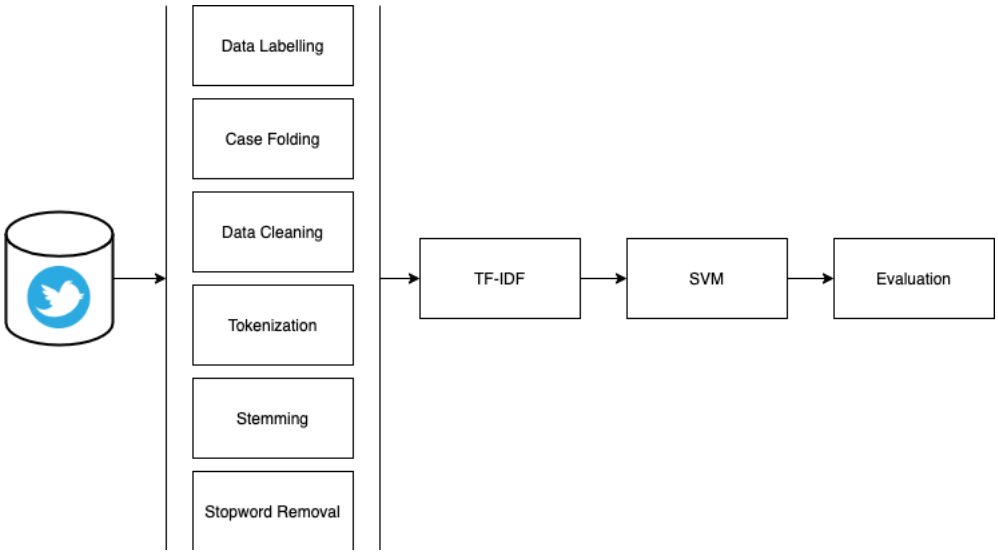

Figure 2. Research Methodology

The first stage for conducting the sentiment analysis process is data collection. In this study, data was taken from Twitter. Retrieval of data from Twitter is quite easy to do because Twitter already provides an Application Programming Interface (API) which is aimed at system developers to make it easier to retrieve data from Twitter.

Apart from Twitter, data can be collected from other sources, such as news portals, online forums, other social media and personal blog sites. The thing that distinguishes Twitter is the way it is collected because not all sites provide APIs that can be used by just anyone.

After the data has been successfully collected into a dataset, the next step is labeling. Labeling here is intended to divide the data into several sentiment classes that will be used in the research. The number of sentiment classes that are widely used is two and three classes, namely negative, neutral and positive. The purpose of this labeling process is to divide the dataset into 2 parts, namely training data and testing data. Training data is data used to train the system to be able to recognize the pattern that is being sought, while testing data is data used to test the results of the training that has been carried out.

Before the dataset is processed to the next stage, the dataset must be tokenized into per sentence (sentence splitting). The results of the tokenized dataset are then stored in .xlsx form. After that, the dataset in this study must first be annotated. The labeling carried out divides comments into sentiments into categories, namely negative and positive. Positive sentiment is given a value of 1 and negative sentiment is given a value of 0. The dataset that has been tokenized based on the sentence is then annotated by 5 annotators. The labeling results carried out by each annotator are then compared with each other and tested to ensure a suitable sentiment value for the comment.

Preprocessing of twitter comment data needs to be done before the classification process so that the dimensions of the vector space model are smaller. By reducing the dimensions of the vector space model the classification process will be faster. The purpose of this pre-processing is to eliminate words that are not suitable for research, homogenize the form of words and reduce the volume of words. The preprocessing stage in this study consisted of several processes, namely data cleaning, case folding, tokenization, stop words removal, and stemming.

1. Data Cleaning: At this stage, the sentences in the dataset are cleaned of anything that can affect the results of the analysis, such as words with two or more repeated characters, links, usernames (@username), hashtags (#), numbers, symbols, excess spaces. , punctuation marks, and numbers. To perform data cleaning, the writer uses a regular expression to match the one to be deleted.

2. Case folding: case folding is done by making all uppercase letters in the dataset lowercase. This stage is carried out so that all characters in the dataset are the same, namely using lowercase letters. By making all words lowercase it will be very helpful to make generalizations [25].

3. Tokenization: Tokenization is a process carried out to break down sentences into pieces of words, punctuation marks, and other meaningful expressions in accordance with the provisions of the language used. In this process, the researcher uses the word_tokenize function provided by the NLTK library.

4. Stopwords Removal: Stopwords Removal is a process done to remove meaningless words. This stage will use the stopwords library. Stopwords add a data dimension to the classification process. The words contained in the stopword list will be removed.

5. Stemming: Stemming is a process carried out to change words that have affixes into their root form by removing affixes such as prefixes, suffixes, and confixes. At this stage, stemming is done using the NLTK library. Stemming is the core of natural language processing techniques for information retrieval which is effective and efficient, and is widely accepted by users. Stemming can also be used to support categorization / classification and clustering processes. Stemming is used to change the form of a word into the root word of the word in accordance with a good and correct morphological structure.

4. RESULTS AND DISCUSSION

When working with datasets, machine learning algorithms work in two stages. Typically the data is split about 80%-20% between the training and testing phases. In supervised learning, dividing the dataset into training data is important for carrying out the learning phase which will later be used as a prediction reference by the algorithm. This process is known as data split.

4.1. Dataset

The dataset used in this study is an opinion in English regarding the iPhone release originating from Twitter. The dataset is in the form of a twitter opinion which is basically a public expression of something through social media. On social media, Twitter itself is more often used in informal language, using grammar that tends to be free, and not a few non-standard words appear in tweets.

All 2000 data documents that have been collected will be processed through the pre-processing stage. Furthermore, the data will be divided into training data and test data with a ratio of 80:20. The training data will be made a classifier model using a predetermined algorithm, while the test data will be predicted using the model that has been built. The accuracy and processing time of the model will be calculated based on the prediction results. This research uses hardware and software with the following specifications (Table 1).

Table 1. Hardware and software specifications

Hardware/Software	Specification
Operating System	macOS Mojave
System Type	64 bit
Processor	Intel® Core® i5-7200U @2.50 GHz (4 CPUs)
RAM	8 Giga Byte
Harddisk	512 Giga Byte
Software	Anaconda (Python 3.4)

After collecting data, data processing or preprocessing is carried out, this stage includes data building activities and continued data cleaning activities so that management can be carried out at the next stage. The preprocessing stage in this study consisted of several processes, namely case folding, data cleaning, tokenization, stop words removal, and stemming.

4.2. Results

The results of sentiment analysis using the SVM algorithm with the use of 1002 data divided into 801 training data and 201 testing data evaluated by calculating the accuracy, precision, recall and F1 score values. Figure 3 below shows the accuracy results using cross validation.

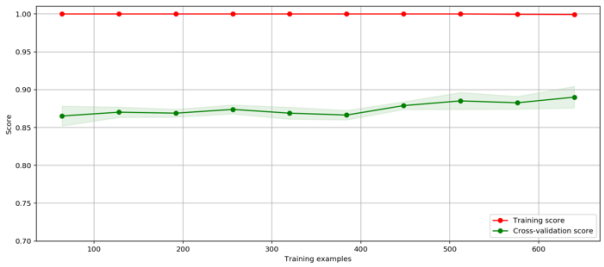

Figure 3. Cross Validation

From the calculation using the cross validation above, the average results for accuracy, precision, recall and F1 score values are obtained, as shown in Table 2.

Table 2. The matrix evaluation results

Matrix Evaluation	Score
Accuracy	89.21%
Precision	92.43%
Recall	95.53%
F1 score	93.95%

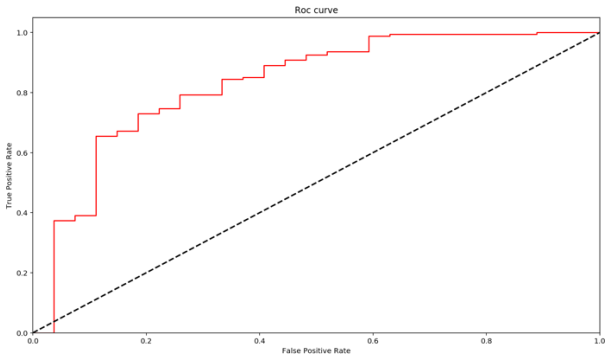

Figure 4. ROC Curve

Area Under Curve (AUC) is the area under the Receiver Operating Characteristic (ROC) curve. If the value is close to one, the model obtained is more accurate. Figure 4 is the ROC curve which has an AUC area of 82.97%. This means that the best model is based on the highest ROC curve at 82.97% with linear kernel parameters. The model obtained is better at 82.97%, because based on the ROC curve of the above category is above 50%. The highest model accuracy based on the resulting ROC curve is 82.97%, meaning that the resulting model has an accuracy of up to 82.97% using the support vector machine algorithm.

5. CONCLUSION

Based on the research that has been carried out on sentiment analysis on the release of the iPhone using the Support Vector Machine (SVM), the conclusion is that, by scraping data, sentiment analysis can be done automatically and quickly. Meanwhile, the preprocessing comments proved to be effective in producing sentences that were important to the sentiment analysis process. And like elaboration on the results and discussion of experiments that have been carried out,, the conclusion that can be drawn from this study is that the SVM classification method has an accuracy of 89.21%. The SVM is widely used as a baseline in text-related tasks with satisfactory results, on several evaluation matrices such as accuracy, precision, recall, and F1 score yielding 89.21%, 92.43%, 95.53%, and 93.95, respectively.

REFERENCES

[1] C. Troussas, M. Virvou, K. J. Espinosa, K. Llaguno, and J. Caro, "Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning," in *IISA 2013*, 2013, pp. 1–6.

[2] B. Liu, "Sentiment analysis and opinion mining," *Synth. Lect. Hum. Lang. Technol.*, vol. 5, no. 1, pp. 1–167, 2012.
[3] T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in *Proceedings of the 2nd international conference on Knowledge capture*, 2003, pp. 70–77.
[4] R. K. Bakshi, N. Kaur, R. Kaur, and G. Kaur, "Opinion mining and sentiment analysis," in *2016 3rd international conference on computing for sustainable global development (INDIACom)*, 2016, pp. 452–455.
[5] E. Sutoyo and A. Almaarif, "Twitter sentiment analysis of the relocation of Indonesia's capital city," *Bull. Electr. Eng. Informatics*, vol. 9, no. 4, pp. 1620–1630, 2020.
[6] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. J. Passonneau, "Sentiment analysis of twitter data," in *Proceedings of the workshop on language in social media (LSM 2011)*, 2011, pp. 30–38.
[7] R. Novendri, A. S. Callista, D. N. Pratama, and C. E. Puspita, "Sentiment Analysis of YouTube Movie Trailer Comments Using Naïve Bayes," *Bull. Comput. Sci. Electr. Eng.*, vol. 1, no. 1, pp. 26–32, 2020, doi: 10.25008/bcsee.v1i1.5.
[8] A. Valdivia, M. V. Luzón, and F. Herrera, "Sentiment analysis in tripadvisor," *IEEE Intell. Syst.*, vol. 32, no. 4, pp. 72–77, 2017.
[9] A. Balahur *et al.*, "Sentiment analysis in the news," *arXiv Prepr. arXiv1309.6202*, 2013.
[10] H. Bhuiyan, J. Ara, R. Bardhan, and M. R. Islam, "Retrieving YouTube video by sentiment analysis on user comment," in *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, 2017, pp. 474–478.
[11] F. Giummolè, S. Orlando, and G. Tolomei, "Trending topics on Twitter improve the prediction of Google hot queries," in *2013 International Conference on Social Computing*, 2013, pp. 39–44.
[12] K. Dey, R. Shrivastava, and S. Kaushik, "Topical stance detection for Twitter: A two-phase LSTM model using attention," in *European Conference on Information Retrieval*, 2018, pp. 529–536.
[13] A. H. Ahmed Abbasi and M. Dhar, "Benchmarking twitter sentiment analysis tools," in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, Iceland, may. European Language Resources Association (ELRA)*, 2014.
[14] S. Doan, B.-K. H. Vo, and N. Collier, "An analysis of Twitter messages in the 2011 Tohoku Earthquake," in *International conference on electronic healthcare*, 2011, pp. 58–66.
[15] E. Haddi, X. Liu, and Y. Shi, "The role of text pre-processing in sentiment analysis," *Procedia Comput. Sci.*, vol. 17, pp. 26–32, 2013.
[16] R. Feldman, J. Sanger, and others, *The text mining handbook: advanced approaches in analyzing unstructured data*. Cambridge university press, 2007.
[17] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science (80-.).*, vol. 349, no. 6245, pp. 255–260, 2015.
[18] I. H. Witten, E. Frank, and M. a Hall, *Data Mining: Practical Machine Learning Tools and Techniques (Google eBook)*. 2011.
[19] A. Aizawa, "An information-theoretic perspective of tf--idf measures," *Inf. Process. \& Manag.*, vol. 39, no. 1, pp. 45–65, 2003.
[20] J. Ramos and others, "Using tf-idf to determine word relevance in document queries," in *Proceedings of the first instructional conference on machine learning*, 2003, vol. 242, no. 1, pp. 29–48.
[21] L. Wang, *Support vector machines: theory and applications*, vol. 177. Springer Science \& Business Media, 2005.
[22] S. Visa, B. Ramsay, A. L. Ralescu, and E. Van Der Knaap, "Confusion Matrix-based Feature Selection.," *MAICS*, vol. 710, pp. 120–127, 2011.
[23] W. Thuiller, M. B. Araújo, and S. Lavorel, "Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales," *J. Veg. Sci.*, vol. 14, no. 5, pp. 669–680, 2003.
[24] F. Gorunescu, "Classification performance evaluation," in *Data Mining*, 2011, pp. 319–330.
[25] D. Jurafsky and J. H. Martin, "Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition." .