Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 6 No. 1 (2022) 177 - 182 ISSN Media Electronic: 2580-0760

The Sentiment Analysis of Spider-Man: No Way Home Film Based on IMDb Reviews

Tb Dzulfiqar Alhafidh[1], P. H. Gunawan[2*], Bambang Ari Wahyudi[3]
School of Computing, Telkom University
[1]tbdzulfiqar@student.telkomuniversity.ac.id, [2]phgunawan@telkomuniversity.ac.id, [3]bambangari@telkomuniversity.ac.id

Abstract

Sentiment analysis is used to determine the overall sentiment in a movie review. The goal of this paper is to investigate the sentiment analysis using multiple classification methods from Spider-Man: No Way Home movie reviews. The review dataset is procured from the IMDb website. Preprocessing methods are used and compared to determine the difference in accuracy score. The methods proposed for this study include Naïve-Bayes, Support Vector Machine (SVM), Stochastic Gradient Descent (SGD), and Decision Tree to find the best accuracy possible. The sentiment analysis of the movie review resulted in 94 positive reviews and 65 negative reviews. The highest accuracy and f1 score for this study are obtained from the SVM and the SGD classifier with an accuracy of 82% and an F1 score of 81% respectively.

Keywords: sentiment analysis, movie reviews, spider-man, classification, IMDb

1. Introduction

The Marvel Cinematic Universe (MCU) franchise has become a global sensation. A series of movies created by Marvel Studios has touched the hearts of millions of people worldwide. Die-hard fans are avidly waiting for the next movie to expand the interconnected fictional world of the MCU. One of the latest anticipated MCU movies released at the end of 2021 is Spider-Man: No Way Home. This movie sets to end the MCU's Spider-Man movie trilogy. It has garnered high praise from the fans who watched it. Many people have reviewed it to be higher than other MCU movies, as shown in Figure 1.

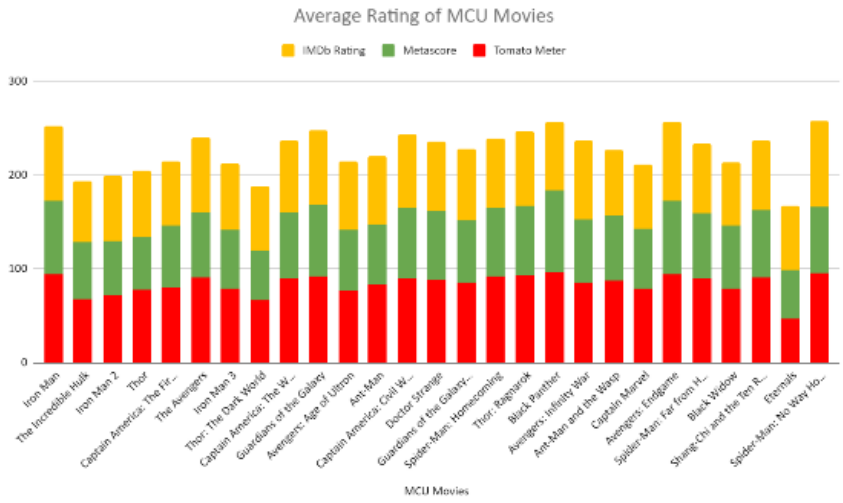

Figure 1. Average Movie Ratings of the current MCU films

Accepted: 03-02-2022 | Received in revised: 25-02-2022 | Published: 28-02-2022

When people review a movie, it isn't always classified as a perfect 10 or a solid 1. It mostly lands between the two numbers as they describe what they like or dislike about said movie. It is then read by other people and affects their judgment on watching the movie, based on the positive and negative sentiments. This is where the study of sentiment analysis comes in. Sentiment analysis is the process of determining a text-based dataset to be positive or negative [1]– [4]. By conducting sentiment analysis on a movie review, people will have an easier time understanding the review's overall sentiment and movie studios will receive the sentiments as feedback to improve their future movies and shows.

There are multiple classification methods in sentiment analysis. One of the more popular methods is using the Naïve-Bayes classifier. Researchers have used this method to analyze movie reviews [5]. Others have improved their methods by adding complex pre-processing methods to increase accuracy [1], [2]. This method is also applicable in product reviews [6]. The other popular method is the Support Vector Machine (SVM) classifier. This method has been proved compatible with product reviews [7], [8]. One research uses frequency-inverse with SVM classifier on movie reviews [3] and another compares it with Naïve-Bayes and random forest method [9]. Other methods used in sentiment analysis are using Decision Tree and Stochastic Gradient Descent classifier. Applications of the decision tree have been used to analyze anti-LGBT tweets on Twitter [10], e-commerce reviews on Google Play Store [11], and Twitter [12]. Applications of stochastic gradient descent are on covid-19 tweets [13], e-commerce tweets [14], and global terrorist attack incident reports [15].

The purpose of this research is to perform a sentiment analysis on the reviews of Spider-Man: No Way Home with the discussed methods above. Later, this work will present a comparison between the four different approaches to classification methods. This comparison will show the ability of several methods (in probability and matrix-vector based) in the sentiment classification.

2. Research Method

2.1 Data Acquisition

The data used for this research was gathered from the Spider-Man: No Way Home movie page on the IMDb website. The reason for choosing IMDb over others is due to the site's ease of use. The data found on the website consisted of the user rating page and user reviews page. The rating used on the website ranged between 1 to 10, 1 being a negative score and 10 being a positive score. This rating system is used in both user rating page and user review page. Figure 2 contains the number of rating votes from the user rating page. The user rating page contains only ratings from the user, without giving a detailed review or feedback. From the data, it tells that most users gave the movie very high ratings.

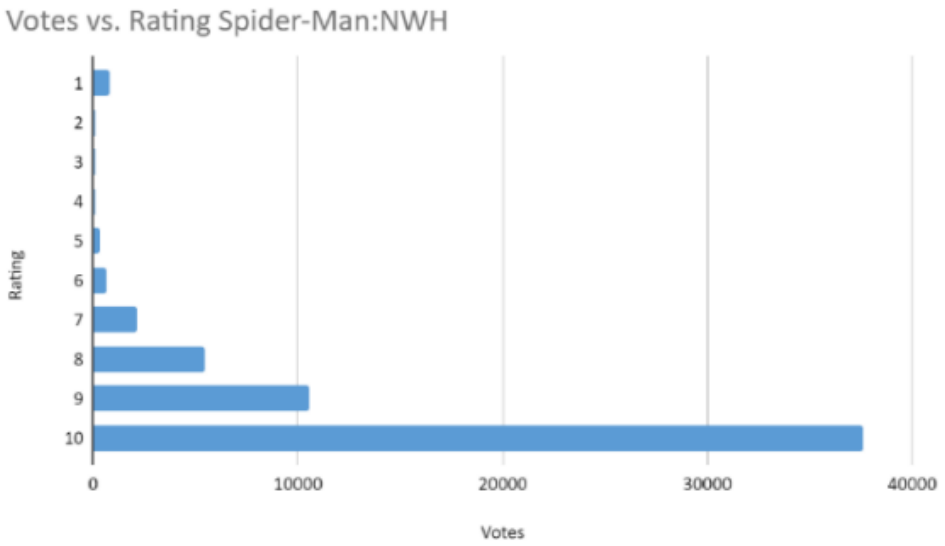

Figure 2. Voters vs Ratings for Spider-Man: NWH User Ratings data

DOI: https://doi.org/10.29207/resti.v6i1.3851

Figure 3 contains the demographic of the users who rate the movie. According to Figure 3, most of the voters are males. Among the male voters, most of them are ranged between 18 to 29 years old. However, the demographic shows a lacking number of female voters compared to male voters. All of the MCU movies up to this point are action superhero movies. This genre resonates with the male audience better than the female audience better due to their high action scene.

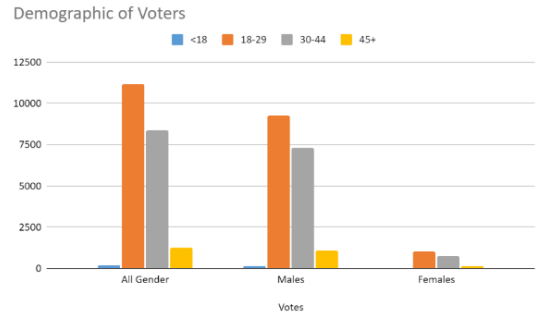

Figure 3. Voters vs Ratings for Spider-Man: NWH User Ratings data

Due to the user ratings, data will not be adequate for the research, the review data will be selected for this research. 200 reviews have been gathered from Spider-Man: No Way Home IMDb page manually. The data is spread in the normal distribution to produce better accuracy in classification. Figure 4 shows the number of votes on each rating in Spider-Man: No Way Home user review data.

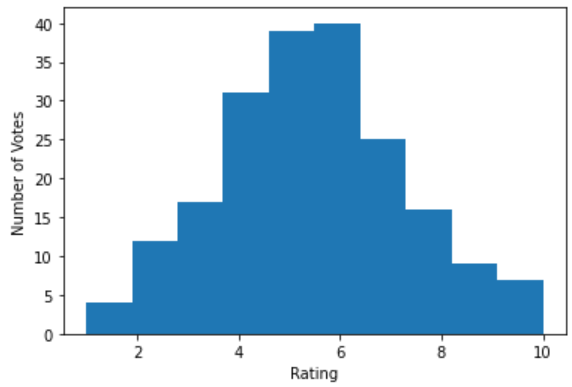

Figure 4. Voters vs Ratings for Spider-Man: NWH User Review data

2.2 Data Pre-processing

Preprocessing data for the research consists of three steps, which include

a) Lowercasing is the process of transforming the text's uppercase letters into lowercase letters. An example of this process is changing the word "REVIEW" into "review".
b) Removing stop words is the process of removing words that is not useful in sentiment analysis. An example of this process is changing the sentence "This movie is bad" to "movie bad" as "This" and "is" are considered as stop words.
c) Lemmatization is the process of converting inflicted words into their base form. This process is preferred over the stemming method due to outputting their base form with meaning. The example of this process is converting the word "am", "are" and "is" into "be".

Classification Methods

The classification methods used in this research are 4 methods; Naïve-Bayes, Support Vector Machine, Decision Tree and Stochastic Gradient Descent.

3.1 Naïve-Bayes

The Naïve-Bayes (NB) classifier is a probabilistic classifier that applies the Bayesian theorem with an assumption of independence [1]. The Bayesian theorem is a theorem used to calculate the posterior probability. The base formula is shown below;

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \quad (1)$$

For $P(c)$ is the prior probability of class c, $P(x)$ is the prior probability of predictor x, $P(c|x)$ is the posterior probability of class c given predictor x and $P(x|c)$is the likelihood probability of predictor x given class c.

Support Vector Machine

The Support Vector Machine (SVM) classifier is a supervised learning model that was introduced by Vapnik in 1992 [11]. The classifier represents classes as hyperplanes in a multidimensional space and tries to locate the maximum marginal space between them. Essentially, it searches for the best possible boundary between positive and negative classes [9].

Decision Tree

Decision Tree is a learning model that uses trees to predict an outcome of an instance [9]. It works by looking at the attributes of the instances and splitting them into different nodes. It can then predict the next node based on the attributes it holds. Variants of the decision tree are found based on their criterion for prediction; Gini Index and Information Gain, with Gini being the more popular. The base formula is shown below.

$$Gini = 1 - \sum_{i=1} (P_i)^2 \quad (2)$$

where P_i denotes the probability of an element classified for a distinct class.

Stochastic Gradient Descent

Stochastic Gradient Descent (SGD) is an iterative learning method that focuses on locating functionality points that can be derived [13]. Stochastic Gradient

DOI: https://doi.org/10.29207/resti.v6i1.3851

Descent works by minimalizing the loss function with the linear function. This method will approach the correct gradient as it considers every sample in a given time, and simultaneously updates the model based on the loss function gradient [14].

General research design of this research can be seen in Figure 5.

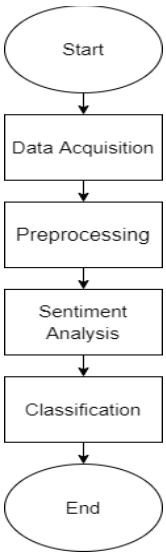

Figure 5. General Research Design

This research is divided into multiple stages, and it is done sequentially. The research stages are divided into 4 stages; Data Acquisition, Pre-Processing, Sentiment Analysis and Classification. The acquisition stage has been explained in this chapter. Therefore, it is not discussed in following chapter.

3. Results and Discussions

3.1 Preprocessing

As previously stated, the data used for this research is the user review dataset. The dataset has 2 parts, which are review title and review comment. Both datasets are preprocessed using the techniques above. The datasets are pre-processed differently, as the review title data has less words to preprocess.

Additional preprocessing techniques is then performed on both datasets. The review title is then removed of its punctuation, numbers, emojis. The review titles are not preprocessed completely because some of them contains only one word, only numbers or all of the words being made up of stop words.

The review comment dataset is also removed of its punctuation, numbers, emojis, and emoticons. Then, we counted the data for the 10 most common words and 10 rare words and then remove them due to the words does not contribute to the sentiment analysis process.

3.2 Sentiment Analysis

As previously discussed, the dataset used for this research is the Spider-Man: No Way Home user reviews on the IMDb website. 200 reviews have been amassed in the normal distribution chart. The reviews are composed of a title and the review comment. After both components have been pre-processed, it is then categorized as positive and negative review by an automatic method.

Figure 6. Result of Sentiment Analysis of the Review Comment

As shown in Figure 6, the result of the sentiment analysis is 94 reviews were defined as positive and 65 reviews were defined as negative. 41 reviews were removed due to assigning the 5-star reviews as the middle point. After the data has been successfully categorized, word clouds are generated from the positive and negative sentiments.

Figure 7. Positive Sentiments Word Cloud

Figure 8. Negative Sentiments Word Cloud

In Figure 7 and 8, two Word louds had generated in order to determine the highest word occurrence in either

DOI: https://doi.org/10.29207/resti.v6i1.3851

sentiment. The bigger the word is in the word cloud, the more often the word is found inside the review comment. The biggest words found in the positive word cloud are "good"," well"," story"," villain", "fan", and "peter". Meanwhile, the biggest words found in the negative word cloud are "villain", "story", "bad", "much" and "plot". The recurrence of the same words in different word clouds shows that both reviews have differing opinions regarding those words.

3.3 Classification

The classification step is where we took the user review dataset that has been categorized and apply different classifiers to find the best accuracy between them. The data used for this classification process is the original title and review comment with the pre-processed title and review comment. This is done to determine whether pre-processing the data affect the accuracy of the classification. The step is then performed ten times for each dataset and classifier. The highest accuracy and f1 score of each iteration are then recorded. This step is performed in Python with the sklearn library. All classifier methods are using their base parameters in their sklearn functions for uniformity and producing better results. When modifying their parameters and kernels for each classifier, it resulted in their accuracy and F1 score drop significantly.

Table 1. Accuracy and F1 score for Original Title Classification

	Original Title	
	Accuracy	F1 Score
Naïve-Bayes	67%	66%
SVM	79%	75.5%
SGD	79%	75.5%
Decision Tree	73%	72%

Table 1 shows the accuracy score and f1 score of the original review title dataset. For the Naïve-Bayes method, it obtained 67% accuracy and 66% f1 score. Then for the SVM method, it obtained 79% accuracy and a 75.5% f1 score. After that for SGD, it obtained 79% accuracy and a 75.5% f1 score. Then for the Decision Tree method, it obtained 73% accuracy and 72% f1 score. The highest accuracy and f1 score overall from this dataset are from the SVM and SGD classifiers.

Table 2. Accuracy and F1 score for Processed Title Classification

	Processed Title	
	Accuracy	F1 Score
Naïve-Bayes	82%	81%
SVM	76%	70%
SGD	61%	75%
Decision Tree	61%	73%

Table 2 shows the accuracy score and f1 score of the preprocessed review title dataset. For the Naïve-Bayes method, it obtained 82% accuracy and 81% f1 score. Then for the SVM method, it obtained 76% accuracy and 70% f1 score. After that for SGD, it obtained 61% accuracy and 75% f1 score. Then for the Decision Tree method, it obtained 61% accuracy and 73% f1 score. The highest accuracy and f1 score overall from this dataset are from the Naïve-Bayes classifier.

Table 3. Accuracy and F1 score for Original Comment Classification

	Original Comment	
	Accuracy	F1 Score
Naïve-Bayes	79%	66.5%
SVM	82%	81%
SGD	73%	68.5%
Decision Tree	76%	75.5%

Table 3 shows the accuracy score and f1 score of the original review comment dataset. For the Naïve-Bayes method, it obtained 79% accuracy and 66.5% f1 score. Then for the SVM method, it obtained 82% accuracy and 81% f1 score. After that for SGD, it obtained 73% accuracy and a 68.5% f1 score. Then for the Decision Tree method, it obtained 76% accuracy and a 75.5% f1 score. The highest accuracy and f1 score overall from this dataset are from the SVM classifier.

Table 4. Accuracy and F1 score for Processed Comment Classification

	Processed Comment	
	Accuracy	F1 Score
Naïve-Bayes	73%	75%
SVM	70%	69.5%
SGD	82%	77%
Decision Tree	75%	74%

Table 4 shows the accuracy score and f1 score of the preprocessed review comment dataset. For the Naïve-Bayes method, it obtained 73% accuracy and 75% f1 score. Then for the SVM method, it obtained 70% accuracy and a 69.5% f1 score. After that for SGD, it obtained 82% accuracy and 77% f1 score. Then for the Decision Tree method, it obtained 75% accuracy and 74% f1 score. The highest accuracy and f1 score overall from this dataset are from the SGD classifier.

6. Conclusion

The sentiment analysis on Spider-Man: No Way movie review shows 94 positive reviews and 65 negative reviews. It is certain that the movie has been greatly enjoyed by reviewers and the audience alike. From the word cloud analysis, it produced similar yet different results. The positive word cloud shows the words “good”, “well”, “peter”, “villain”, and “story” as their most occurring words. Meanwhile, the negative word cloud shows the words “bad”, “story”, “villain”, “plot” and “time” as their most occurring words. The words “villain” and “story” appearing in both word clouds are proof that the reviews in both sentiments have conflicting opinions when using those words. From the classification experiment, the pre-processing step is essential for improving accuracy. Naive-Bayes is the most accurate classifier in the processed title dataset. SVM is the most accurate classifier in the original title and original comment dataset. SGD is the most accurate classifier in the processed title and processed comment

DOI: https://doi.org/10.29207/resti.v6i1.3851

dataset. Overall, the best classifiers for the dataset are the SVM and SGD methods. The highest accuracy and f1 score for this study are obtained from the SVM and the SGD classifier with an accuracy of 82% and an F1 score of 81% respectively. For future works, we could use a larger dataset with balanced positive and negative reviews and modify the model and algorithm to achieve an even higher accuracy score.

Reference

[1] R. Bintang Purnomoputra and U. Novia Wisesty, "Sentiment Analysis of Movie Reviews using Naïve Bayes Method with Gini Index Feature Selection," OPEN ACCESS J DATA SCI APPL, vol. 2, no. 2, pp. 85–094, 2019, doi: 10.34818/JDSA.2019.2.36.

[2] M. B. Hamzah, "Classification of Movie Review Sentiment Analysis Using Chi-Square and Multinomial Naïve Bayes with Adaptive Boosting," Journal of Advances in Information Systems and Technology, vol. 3, no. 1, 2021, [Online]. Available: https://journal.unnes.ac.id/sju/index.php/jaist

[3] M. A. Muslim, "9. Improve the Accuracy of Support Vector Machine Using Chi Square Statistic and Term Frequency Inverse Document Frequency on Movie Review Sentiment Analysis."

[4] R. Maulana, P. A. Rahayuningsih, W. Irmayani, D. Saputra, and W. E. Jayanti, "Improved Accuracy of Sentiment Analysis Movie Review Using Support Vector Machine Based Information Gain," in Journal of Physics: Conference Series, Nov. 2020, vol. 1641, no. 1. doi: 10.1088/1742-6596/1641/1/012060.

[5] Y. Nurdiansyah, S. Bukhori, and R. Hidayat, "Sentiment analysis system for movie review in Bahasa Indonesia using naive bayes classifier method," in Journal of Physics: Conference Series, Apr. 2018, vol. 1008, no. 1. doi: 10.1088/1742-6596/1008/1/012011.

[6] T. Hariguna, W. Maulana Baihaqi, and A. Nurwanti, "Sentiment Analysis of Product Reviews as A Customer Recommendation Using the Naive Bayes Classifier Algorithm," International Journal of Informatics and Information Systems, vol. 2, no. 2, pp. 48–55, 2019.

[7] E. Tyagi and A. K. Sharma, "Sentiment Analysis of Product Reviews using Support Vector Machine Learning Algorithm," Indian Journal of Science and Technology, vol. 10, no. 35, pp. 1–9, Jun. 2017, doi: 10.17485/ijst/2017/v10i35/118965.

[8] S. Dey, S. Wasif, D. S. Tonmoy, S. Sultana, J. Sarkar, and M. Dey, "A Comparative Study of Support Vector Machine and Naive Bayes Classifier for Sentiment Analysis on Amazon Product Reviews," in 2020 International Conference on Contemporary Computing and Applications, IC3A 2020, Feb. 2020, pp. 217–220. doi: 10.1109/IC3A48958.2020.233300.

[9] M. Guia, R. R. Silva, and J. Bernardino, "Comparison of Naive Bayes, support vector machine, decision trees and random forest on sentiment analysis," in IC3K 2019 - Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 2019, vol. 1, pp. 525–531. doi: 10.5220/0008364105250531.

[10] V. A. Fitri, R. Andreswari, and M. A. Hasibuan, "Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm," in Procedia Computer Science, 2019, vol. 161, pp. 765–772. doi: 10.1016/j.procs.2019.11.181.

[11] Nurfaizah, T. Hariguna, and Y. I. Romadon, "The accuracy comparison of vector support machine and decision tree methods in sentiment analysis," in Journal of Physics: Conference Series, Nov. 2019, vol. 1367, no. 1. doi: 10.1088/1742-6596/1367/1/012025.

[12] A. Bayhaqy, S. Sfenrianto, K. Nainggolan, and E. R. Kaburuan, "Sentiment Analysis about E-Commerce from Tweets Using Decision Tree, K-Nearest Neighbor, and Naïve Bayes," Jul. 2018. doi: 10.1109/ICOT.2018.8705796.

[13] V. Dwi Antonio, S. Efendi, and H. Mawengkang, "Sentiment analysis for covid-19 in Indonesia on Twitter with TF-IDF featured extraction and stochastic gradient descent," Int. J. Nonlinear Anal. Appl, vol. 13, no. 1, pp. 2008–6822, 2022, doi: 10.22075/ijnaa.2021.5735.

[14] N. Khotimah, M. Yamin Darsyah, and I. M. Nur, "Analisis Sentimen Terhadap Review E-Commerce Dengan Metode Stochastic Gradient Descent." [Online]. Available: http://repository.unimus.ac.id

[15] Shadi. Diab, "Optimizing stochastic gradient descent in text classification based on fine-tuning hyper-parameters approach. a case study on automatic classification of global terrorist attacks." Accessed: Jan. 25, 2022. [Online]. Available: arXiv preprint arXiv:1902.06542 (2019).---

DOI: https://doi.org/10.29207/resti.v6i1.3851