

## RESEARCH ARTICLE

 OPEN ACCESS

Received: 21.07.2021

Accepted: 20.11.2021

Published: 24.12.2021

**Citation:** Almasani SAM, Qaid WAA, Saif JAM, Alqubati IAA (2021) Fuzzy rule based sentiment analysis for finding University Student Satisfaction in Yemen. Indian Journal of Science and Technology 14(44): 3264-3269. <https://doi.org/10.17485/IJST/v14i44.1283>

\* **Corresponding author.**

[dr.wadqaid@gmail.com](mailto:dr.wadqaid@gmail.com)

**Funding:** None

**Competing Interests:** None

**Copyright:** © 2021 Almasani et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](#))

**ISSN**

Print: 0974-6846

Electronic: 0974-5645

# Fuzzy rule based sentiment analysis for finding University Student Satisfaction in Yemen

**Siham Abdulmalik Mohammed Almasani<sup>1\*</sup>, Wadea Ahmed Abdo Qaid<sup>2,1</sup>, Jamil Abdulhamid Mohammed Saif<sup>3</sup>, Ibrahim A A Alqubati<sup>4</sup>**

**1** Assistant professor, Faculty of Computer Science and Engineering, Hodeidah University, Yemen

**2** Dean Faculty of Engineering and Technology, Future University, Yemen

**3** Department of Computer and Information Systems, College of Community, University of Bisha, KSA

**4** Computer Department, Community College, Najran University, KSA

## Abstract

**Objectives:** To analyze the sentiment and understand Student's opinions and satisfaction about university, and the services which introduce by it to reach accuracy in evaluation of the university. **Methods:** The sentiment of students was analyzed and quantified the satisfaction towards the University by building an algorithm using a set of fuzzy rules. This study proposes an integrated optimization method: applying a set of fuzzy rules with two lexicons (SentiWordNet and AFINN lexicon) besides to other parameter. The lexicons in sentiment analysis are individual words that can be considered as a unit of opinion information. The proposed fuzzy system integrates Natural Language Processing techniques to analyze and quantify student's satisfaction towards the University by classify the comments into very positive, positive, very negative, negative, or neutral sentiment classes. **Findings:** with this approach, the accuracy is more 0.891 compared to Support Vector Machines (SVM), Naïve Bayes, The Fuzzy Inference to analyze the sentiment with their own lexicon Opinion Words Lexicon and classified the sentiment into 2 class-positive, negative and The Fuzzy Inference to analyze the sentiment with only SentiWordNet lexicon in evaluation obtained for University evaluation, the evaluation has been verified by simulation results on MATLAB. **Novelty:** The novelty of this study lies in the formulation of few fuzzy rules to evaluate the sentiment class of tweets, and the proposed model can be adapted to any lexicon.

**Keywords:** Sentiment analysis; Fuzzy rule; linguistic variables; Students comments analysis

## 1 Introduction

Sentiment Analysis is a difficult problem, where the users can freely express their opinions and feelings on different facility and topics<sup>(1)</sup>. Sentiment Analysis can be

used to improve services. Over the last decades, there has been a lot of papers have been presented in on sentiment analysis to analyzes people's opinions, evaluations, and emotions towards entities such as products, services, individuals, and organizations but there was not enough accuracy for the correct assessment. From example, the traditional techniques of classifying the sentiment are not scalable, which was motivating the researchers to develop automated techniques. Most of the works have classified the social media posts according to the sentiment contained in them. Many papers have used machine learning techniques like Naïve Bayes<sup>(2-4)</sup> and other papers used Support -Vector Machines (SVM) for Sentiment Analysis from tweets<sup>(5,6)</sup> but the accuracy was not high. A supervised machine learning-based sentiment analysis system for analyzing student reviews about teacher's performance is proposed<sup>(7)</sup>. The Support Vector Machine (SVM) is used for classifying reviews into positive, negative, or neutral. The results show the satisfactory performance of the proposed system with respect to comparing methods.

In<sup>(8)</sup> proposed the Fuzzy Inference to analyze the sentiment and classified the sentiment into 2 classes- positive, negative. They have constructed their own lexicon Opinion Words Lexicon by performing some modifications on SentiWordNet data. But this paper<sup>(8)</sup> can not be used with any lexicon: SentiWordNet, AFINN, and VADER. In<sup>(9)</sup> they have used only SentiWordNet lexicon and don't use any other lexicon. The sentiment score for each term in the tweet is computed as the difference between positive and negative scores obtained from the lexicon.

A fuzzy system is a powerful approach for modeling systems, which is introduced by (Zadeh, 1965). The basis of fuzzy logic is to consider the inputs and outputs of the system as fuzzy sets and build the fuzzy rules. Each fuzzy set contains elements that have varying degrees of membership<sup>(10,11)</sup>. Fuzzy rules relate input variables to output variables. These rules represent the expert's knowledge of the system. Mamdani's approach is one of the most common fuzzy inference methods which defines the output of each fuzzy rule as a fuzzy set for the output variable (Mamdani, 1974). Defuzzification is the last step over the system which is a mapping process from a fuzzy space defined of output into a crisp. We have used the concept of fuzzy system to create our system to determine the students's sentiment and quantify students Satisfaction towards the University.

In this paper the following technique was used, firstly, the text was pre-processed to get useful information by removing the unnecessary information. Because the students post additional information when express toward there's sentiment by using abbreviations, URLs, emoticons and hashtags. Secndly, the fuzzy Logic with two lexicans (SentiWordNet and AFINN lexicons beside the confirmation parameter as inputs) to process the students's comments was applied. The used technique was applied in future university in Yemen.

## 2 Materials and M ethods

### 2.1 Proposed Fuzzy System for Sentiment Analysis of Students Towards the University

In this section, the details of the proposed fuzzy logic algorithm were presenting to analyze the sentiment of the Students toward the university. Figure 1 describes the fuzzy logic algorithm. The fuzzy system is most useful in modeling complex systems because it uses linguistic variables to make it easier to find solutions. In our proposed algorithm, proposed the Fuzzy Inference to analyze the sentiment and classified the sentiment into 5 classes (positive, very positive, natural, negative, and very negative). We have used SentiWordNet and AFINN lexicons beside the confirmation parameter as inputs to the fuzzy model. SentiWordNet and AFINN lexicons work excellently well on social media text and it doesn't require any training data.

Based on the polarity of the words, which can be positive or negative, the outputs are assigned either (positive, very positive, natural, negative, and very negative), however, polarities of the output are increased or decreased with the presence of adverbs (In this work, we express about this as input to our proposed system as confirmation parameter), such as very, not, absolutely and never and are reversed with the present of negative words. To build this form, a set of rules is developed to evaluate input and produce proper output. The results show the satisfactory performance of the proposed algorithm and more accuarcy with respect to comparing methods.

### 2.2 Text pre-processing

Users always post information that depicts sentiment, besides that; they post additional information using abbreviations, emoticons, slang, or URLs. Thus, the information which they posted needs to be pre-processed to get useful information by removing the additional data. So, we have eliminated all of the additional data which they don't carry any sentiment.

### 2.3 The used sentiment lexicons

A sentiment lexicon is a list of words which are generally labeled according to their semantic orientation as either positive or negative. In the article, two different sentiment lexicons have been used beside to construct our own Opinion Words Lexicon:

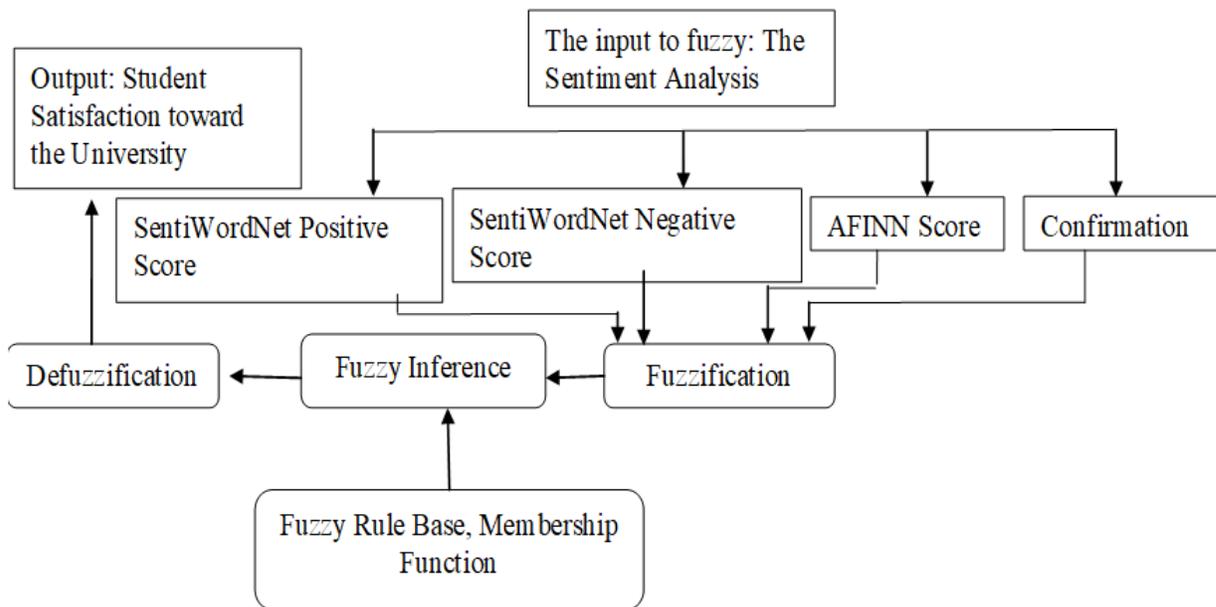


Fig 1. The proposed fuzzy logic algorithm to analyze the students’ sentiment towards the University

SentiWordNet<sup>(12,13)</sup>, AFINN<sup>(14)</sup>. The pre-processed text is used with these two lexicons to compute the positive and negative scores for each post.

In the SentiWordNet three numerical scores relating to positive, negative, and neutral were annotated. Each score ranges from 0 to 1, and their sum is 1. For example, in-text ”I am very impressed from her thinking method”, the word ”very” gives an amount of assurance, and the word ”impressed” is an opinion word. The SentiWordNet method obtains the scores of each word from this lexicon and each word has positive and negative scores these scores are computed as the following:

$$\begin{aligned}
 \text{score}(b) &= \text{score}_{pos}(w), \text{ if } (\text{score}_{pos}(w) > \text{score}_{neg}(w)) \\
 \text{score}(b) &= \text{score}_{neg}(w), \text{ if } (\text{score}_{pos}(w) < \text{score}_{neg}(w))
 \end{aligned}
 \tag{1}$$

The AFINN lexicon is a list of English terms rated for valence with an integer between -5 negative and + 5 positive. The AFINN method fetches the score of each word, if it is greater than 0 it is a positive word, and if less than 0 it is a negative word.

$$\text{score}(b) = \begin{cases} \text{score}_{pos}(w), & \text{if } (\text{score}_{pos}(w) > 0) \\ \text{score}_{neg}(w), & \text{if } (\text{score}_{pos}(w) < 0) \end{cases}
 \tag{2}$$

### 2.4 The Proposed Fuzzy System

To build the fuzzy model for Analyzing Student sentiment the popularly Mamdani fuzzy model was used. The Mamdani style fuzzy inference process is performed in four steps: Fuzzification of input variables, Rule evaluation (inference), Aggregation of the rule outputs, and Defuzzification.

The inputs of the fuzzy model are Confirmation, AFINN score; theSentiWordNetpositive and negative score of each post obtained from the above phase is fuzzified using Pi-shaped membership function. When the Pi-shaped membership function is used, each linguistic term T involves four key points, x, a, b, c, d associated with the change of pattern of the fuzzy membership. The Pi-shaped membership functions of fuzzy model inputs were graphically presented in Figures 2, 3, 4 and 5.

The fuzzy rules for Analyzing Student Satisfaction were described in Table 1. The rules were obtained by the intersection of four input variables (Confirmation, AFINN score, the SentiWordNet positive and negative score), each with three fuzzy subsets Except the Confirmation input with two fuzzy subsets. The rules were devised based on the assumption that the higher score (positive; very positive; negative; very negative or natural) indicates the sentiment. In the case of common scores, the sentiment is neutral.

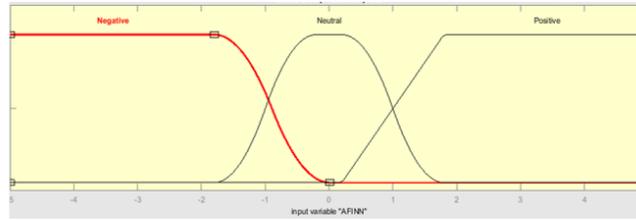


Fig 2. Pi-shaped fuzzy membership sets for AFINN score input

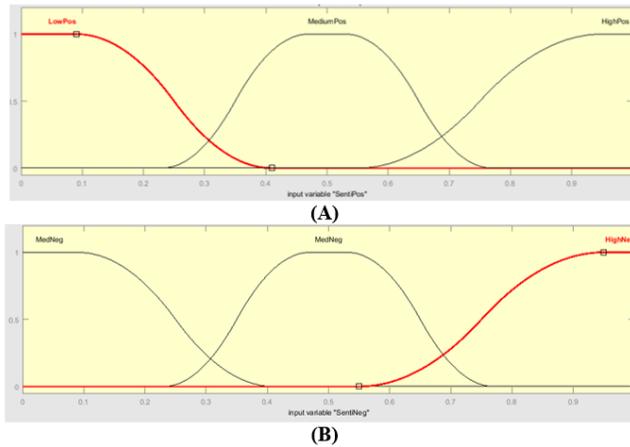


Fig 3. Pi-shaped fuzzy membership sets for the SentiWordNet positive and negative score inputs.

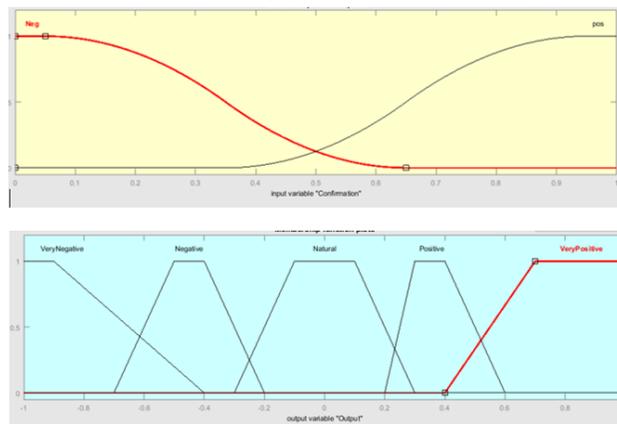


Fig 4. Pi-shaped fuzzy membership sets of output for the sentiment analysis for Analyzing Student Satisfaction toward the university.

Experts consider each variable as a linguistic variable (LV), define a set  $\langle \alpha_i, T(\alpha_i), X_i, G_i, M_i \rangle, i = 1, n$ , where  $\alpha_i$ – name of LV;  $T(\alpha_i)$  - term-set LV  $\alpha_i$ ;  $X_i$  – domain LV  $\alpha_i$ ,  $G_i$  - syntactic rule;  $M_i$  - semantic rule<sup>(14-16)</sup>. For each LV  $\alpha_i$ , experts give the term-set  $T(\alpha_i)$  assets of fuzzy variables

In this work, the selected parameters will be determined at the verbal level by the following linguistic variables (LV):

- LP<sub>1</sub> – the SentiWordNet positive score has base-sets  $T(LP_1)=\{$
- LP<sub>2</sub> – the SentiWordNet negative score has base-sets  $T(LP_2)=\{$
- LP<sub>3</sub> – AFINN score has base-sets  $T(LP_3)=\{$
- LP<sub>4</sub> – Confirmation has base-sets  $T(LP_4)=\{$

Table 1 The fuzzy rules for Analyzing Student Satisfaction.

**Table 1.** The fuzzy rules for Analyzing Student Satisfaction

Rule	AFINN score	SentiWordNetpositive	SentiWordNetnegative	Confirmation	Sentiment
1	Positive	HighPos	LowNeg	Pos	very positive
2	Positive	MediumPos	LowNeg	Pos	very positive
3	Positive	LowPos	LowNeg	Pos	Positive
4	Positive	HighPos	LowNeg	Neg	Negative
5	Negative	LowPos	HighNeg	Pos	VeryNegative
6	Negative	LowPos	MediumNeg	Pos	VeryNegative
7	Natural	LowPos	LowNeg	-	Natural
8	.....	....	....	....	.....

Finally, after the inputs and output are determined, the fuzzy rules are built. The program to analyze the sentiment of Students toward the university is created as shown in Figure 5.

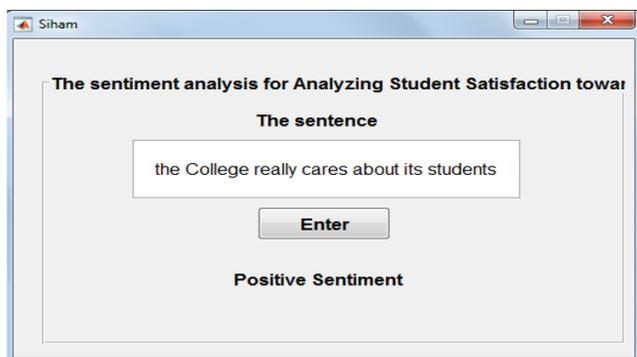


Fig 5. The result of analysis the sentiment of Students toward the university

The user enters the Sentences which express about students’ sentiment toward the university and then the program analyze these sentiments and classify it.

### 3 Results and Discussion

The program results for sentiment analysis should be evaluated by Precision criteria. Therefore, in this case, the application of expert assessments is desirable and gives more reliable estimates. Table 2 shows the expert evaluation for the modeling results using the work program and the similar programs created based on the different models. Thus, from the comparison results, we can conclude that the characteristics of the work algorithm which is developed in the article are better.

Table 2. Assessment of the results using the different types of methods

Methods	Precision
Our work	0.891
Support Vector Machines (SVM)	0.694
Naïve Bayes	0.456
The Fuzzy Inference to analyze the sentiment with their own lexicon Opinion Words Lexicon and classified the sentiment into 2 class- positive, negative	0.721
The Fuzzy Inference to analyze the sentiment with only SentiWordNet lexicon	0.788

As shown in Table 2. The precision of the proposed method was more than other methods.

## 4 Conclusion

This study proposes a fuzzy model for Sentiment Analysis of Student Satisfaction toward the Future University in Yemen. The novelty is the formulation of few fuzzy rules to evaluate the sentiment class of tweets, and the proposed model can be adapted to any lexicon. The proposed fuzzy model was implemented using two different lexicons SentiWordNet, AFINN as inputs beside the confirmation input. Comparison with other methods (Support Vector Machines (SVM), Naïve Bayes, The Fuzzy Inference to analyze the sentiment with their own lexicon and The Fuzzy Inference to analyze the sentiment with only SentiWordNet lexicon) reveals that the proposed fuzzy rule-based model performs consistently the best.

This work employed the fuzzy rules due to the fact that the Fuzzy systems can deal with ambiguity. The fuzzy rules incorporate the fuzziness of positive and negative scores and deal with reasoning and give closer views to the exact sentiment values.

In the future, this work can be implemented in other domains like evaluation the consumer satisfaction with products, etc. this study can extend by incorporating fuzzy inferencing into deep neural network models.

## References

- 1) Birjali M, Kasri M, Beni-Hssane A. A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*. 2021;226:107134. Available from: <https://dx.doi.org/10.1016/j.knosys.2021.107134>.
- 2) Kastrati Z, Dalipi F, Imran AS, Nuci KP, Wani MA. Sentiment Analysis of Students' Feedback with NLP and Deep Learning: A Systematic Mapping Study. *Applied Sciences*. 2021;11(9):3986. Available from: <https://dx.doi.org/10.3390/app11093986>.
- 3) Ho IMK, Cheong KY, Weldon A. Predicting student satisfaction of emergency remote learning in higher education during COVID-19 using machine learning techniques. *PLOS ONE*. 2021;16(4):e0249423. Available from: <https://dx.doi.org/10.1371/journal.pone.0249423>.
- 4) Parveen H, Pandey S. Sentiment analysis on Twitter Data-set using Naive Bayes algorithm. *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*. 2016;p. 416–419. doi:10.1109/ICATCCCT.2016.7912034.
- 5) Hamdan H, Hussam, Béchet, Frederic, Bellot, Patrice. Experiments with DBpedia, WordNet and SentiWordNet as resources for sentiment analysis in micro-blogging. *Second Joint Conference on Lexical and Computational Semantics (SEM)*. 2013;2:455–459. Available from: <https://aclanthology.org/S13-2075.pdf>.
- 6) Windasari IP, Uzzi FN, Satoto KI. Sentiment analysis on Twitter posts: An analysis of positive or negative opinion on GoJek. *2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*. 2017;p. 266–269. Available from: [10.1109/ICITACEE.2017.8257715](https://doi.org/10.1109/ICITACEE.2017.8257715).
- 7) Esparza GG, De-Luna A, Zezzatti AO, Hernandez A, Ponce J, Álvarez M, et al. A Sentiment Analysis Model to Analyze Students Reviews of Teacher Performance Using Support Vector Machines. *International Symposium on Distributed Computing and Artificial Intelligence, Springer, Cham*. 2018;p. 157–164. Available from: [https://link.springer.com/chapter/10.1007/978-3-319-62410-5\\_19](https://link.springer.com/chapter/10.1007/978-3-319-62410-5_19).
- 8) Srivastava R, Bhatia MPS. Quantifying modified opinion strength: A fuzzy inference system for Sentiment Analysis. *2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. 2013;p. 1512–1519. doi:10.1109/ICACCI.2013.6637404.
- 9) Haque A, Rahman T. Sentiment Analysis by Using Fuzzy Logic. *International Journal of Computer Science, Engineering and Information Technology*. 2014;4(1):33–48. Available from: <https://dx.doi.org/10.5121/ijcseit.2014.4104>.
- 10) Zadeh LA. The Concept of a Linguistic Variable and its Application to Approximate Reasoning. *Information Science*. 1974;8:199–249. Available from: <https://www.sciencedirect.com/science/article/abs/pii/0020025575900365>.
- 11) Zadeh LA. Fuzzy logic—a personal perspective. *Fuzzy Sets and Systems*. 2015;281:4–20. Available from: <https://dx.doi.org/10.1016/j.fss.2015.05.009>.
- 12) Stefano B, Andrea E, Fabrizio S. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *European Language Resources Association*. 2010;10:2200–2204. Available from: <https://aclanthology.org/L10-1531/>.
- 13) Ramshankar N, Prathap PMJ. A novel recommendation system enabled by adaptive fuzzy aided sentiment classification for E-commerce sector using black hole-based grey wolf optimization. *Sādhanā*. 2021;46(3):1–24. Available from: <https://dx.doi.org/10.1007/s12046-021-01631-2>.
- 14) Ghiassi M, Skinner J, Zimbra D. Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with Applications*. 2013;40(16):6266–6282. Available from: <https://dx.doi.org/10.1016/j.eswa.2013.05.057>.
- 15) Almasani SAM, Finaev VI, Qaid WAA, Tychinsky AV. The Decision-making Model for the Stock Market under Uncertainty. *International Journal of Electrical and Computer Engineering (IJECE)*. 2017;7(5):2782. Available from: <https://dx.doi.org/10.11591/ijece.v7i5.pp2782-2790>.
- 16) Almasani SAM, Finaev VI, abdo qaid WA, vladimirovich tychinsky A. Assessing the Current State of the Stock Market Under Uncertainty. *Journal of Theoretical and Applied Information Technology*. 2016;89(1). Available from: [http://iues.sfedu.ru/wp-content/uploads/2016/08/Almasani-skopus2\\_2016.pdf](http://iues.sfedu.ru/wp-content/uploads/2016/08/Almasani-skopus2_2016.pdf).