

Students' Sentiment Analysis Using Natural Language Toolkit in Machine Learning for Module Evaluation

Carine Umunyana^{1,*}, Gerard Tuyizere², and Anaclet Mbarushimana²

ABSTRACT

This paper presents a combination of natural language toolkit (NLTK) in machine learning for sentiment analysis used for module evaluation. The module evaluation is typically done at the end of each module. Dataset of 300 students evaluating each module is conducted with excellent, very good, good, fair, and poor sentiments, delivers valuable perceptions into the overall teaching and lecturing quality and decision making for enlightening methodology of teaching and approaches. This paper demonstrates sentiment analysis model trained using logistic regression algorithm in Machine Learning to evaluate the sentiments given by students in their module evaluation. A study comparison has been done between the proposed model and other sentiment analysis for module evaluation. The results of experiments have been analyzed for decision-making.

Keywords: Machine learning and module evaluation, natural language toolkit, sentiment analysis.

Submitted: March 30, 2023

Published: March 03, 2024

 10.24018/lejeng.2024.9.1.3006

¹ Ministry of Environment, Rwanda.

² Rwanda Polytechnic, Integrated Polytechnic Regional College (RP/IPC Ngoma), Rwanda.

* Corresponding Author:
e-mail: umucarin@gmail.com

1. INTRODUCTION

Nowadays, the teaching system embodies a background that is endlessly improved by an enormous number of low facts generated every second in various formats and most often carried useful and valuable information. Discovering and retrieving the unknown "information" from the ocean of educational data establishes one of the most benefits that sentiment analysis can offer.

Sentiments given by students are a valuable foundation of data not only for evaluating a lecture/module but also for reforming strategies and quality for academic improvement. The sentiment analysis from the students' evaluation of a module entity was considered.

Sentiment analysis can be applied to a document, sentence, or word level. Though, because of the high number of documents, manual management of sentiments is unfeasible. Hence, digitalized data manipulation is of the great role. Sentiment from the text-based (sentence or document) using natural language processing (NLP). Most of the research published until 2016-2017 used full NLP techniques.

Natural language processing (NLP) technique is used by Sentiment analysis to define the attitude of datasets. Huge amount of data that is produced today is unstructured, which involves processing to produce information.

Some samples of unstructured data are news on social media post, feedback on some topics and search history. The technique of investigating natural language and retrieving meaningful information falls under the field of Natural Language Processing (NLP). The Natural Language Toolkit (NLTK) is used as a commonly used NLP library in Python, to analyze textual data.

In this paper, a dataset of survey done over 300 students to evaluate the module was conducted for NLP with different data cleaning methods. Once the dataset is ready for processing, the training model on pre-classified datasets about module evaluation and use the model to classify the sample datasets into excellent, very good, good, fair and poor sentiments.

2. MATERIALS AND METHODS

2.1. Sentiment Analysis and Related Work

Sentiment analysis, in general, is an assignment that focuses on discovering and appreciating feelings toward a variable, which can be a topic, event or a person, commonly, the intention of sentiment analysis is to get clear feelings of users, recognize the views they demonstrate, and then locate them into positive, negative, and neutral classes. Natural language and machine learning algorithms are



used by sentiment analysis arrangements to bring light on, fetch, and extract figures and opinions from vast amounts of textual information [1].

Mostly, sentiment investigation can be executed in three different classes: the document, sentence, aspect level. Sentiment analysis at the document level intends to bring out the sentiments of users by exploring the document entirely. Sentence level investigation tends to be simple as the target is to identify the polarity of sentences instead of the entire document. Aspect level sentiment investigation emphasize on identifying aspects or attributes expressed in reviews and on classifying the opinions of users towards these aspects.

Sentiment investigation has been commonly useful in different application arena, especially in business and social networks, for different reason. Some well-known sentiment analysis business applications include product and services reviews [2], financial markets [3], customer relationship management [4], and marketing strategies and research, among others. About social media applications, the most known application of sentiment analysis is to control the aspect of a specific brand on Twitter or Facebook [5], and explore the reaction of people given a crisis; e.g., COVID-19 [6]. Another important application domain is in politics [7], where sentiment investigation can be helpful for the campaigns in election of candidates running for political positions.

Nowadays, sentiment analysis has also involved a great deal in research attention in the applied sciences and technology arena. In contrast to the above-mentioned area of business or social networks, which emphasize a single participant, the investigation on sentiment analysis in the education arena considers multiple shareholders of education, including lecturers/tutorials, students, decision-makers, and institutions. Specifically, sentiment

investigation is mainly applied to improve lecturing, management, and evaluation by analyzing students' attitudes and behavior towards courses, platforms, institutions, and teachers.

Referring to past literature, one study [8] on "sentiment analysis (SA) in the education domain focused on detecting the approaches and resources used in SA and identifying the main benefits of using SA on education data", the study is developed from this paper.

3. RESEARCH DESIGN

3.1. Introduction

As depicted on Fig. 1, the following actions were performed in this architecture:

- Data gathering: The datasets are gathered from three hundred students from different departments evaluating the module/course aspects, and the data are captured on Google Sheets (which was the problem in making the decision from the data) and loaded into the NLTK package for future use.
- Data cleaning: This process includes removing duplicate data, wrong data, empty cells, and data in the wrong format.
- Training and testing model: In this step, the dataset is trained through a Naïve Bayesian classifier.
- Building testing model: Naïve Bayesian Classifier can be used to build the model.

3.2. Performance Measures

There are three commonly used performance measurements: accuracy, sensitivity, and specificity. The accuracy of classifiers is the percentage of correctness of outcome among the test sets exploited in this study. It is defined in

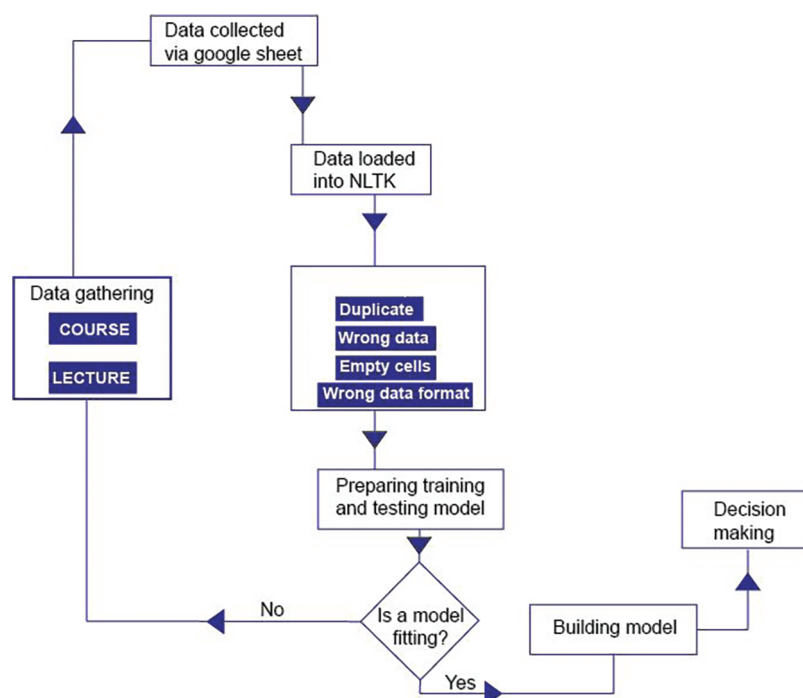


Fig. 1. Design of sentiment analysis process.

(1)–(3). The sensitivity is referred to as the true positive rate, and the specificity is the true negative rate. Both sensitivity and specificity used for measuring the factors that affect the performance are presented in (2) and (3), respectively:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN + FN} \quad (3)$$

where True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN).

4. RESULTS AND TABLES

The training and tests were done using Google Colab. The results presented in the Table I are evaluation captured from students evaluating the course at the end of the whole

TABLE I: REPRESENTATION OF DATA RETRIEVED FROM THE STUDENT WITH THEIR SENTIMENTS

S/N	Name	Rating	Review
0	19RP04203	Module leader's mastery of content	Excellent
1	19RP04692	Quantity and relevance of assessment	Good
2	19RP00165	Quantity and relevance of assessment	Good
3	19RP02249	Linking the content with real life and job situation	Very good
4	19RP09028	Module leader's mastery of content	Excellent
5	19RP10677	Quantity and relevance of assessment	Good
6	19RP08509	Quantity and relevance of assessment	Good
7	18RP09142	Quantity and relevance of assessment	Good
8	19RP07093	Linking the content with real life and job situation	Very good
9	19RP03090	Linking the content with real life and job situation	Very good

lecture/module. This will help the course owner (lecturer, assistant lecturer, or tutorial assistant) get an overview of the progression of the course.

The data that was used in the survey was collected through Google Sheets and received on Google Driver. The following codes have been used to connect to these data:

```
• from google.colab import drive
• drive.mount('/content/drive')
```

Some important libraries have been imported to facilitate the running environment and processing of datasets:

```
• import pandas as pd
• import numpy as np
• import tensorflow as tf
• import sklearn
• import os, sys, math
```

The machine cannot process the language in its original way accurately; the manipulation of the language in a way that machines can understand is needed, giving a sense to the data through the process called tokenization:

```
• from tensorflow.keras.preprocessing.
text import Tokenizer
```

Sentiment analysis is a process of identifying an attitude of the phrase, text or word. Training data set to train a model has been generated. Supervised learning machine learning process has been used, which requires associating each dataset with a “sentiment” for training. In this paper, the model used the “excellent”, “very good”, “good”, “fair” and “poor” sentiments:

```
• df_student_evaluation.head(10)
```

The data are prepared for training the logistic regression classifier class is used:

```
• X, y = df_student_evaluation['review'].
astype('str'), df_student_evaluation
['sentiment']
• X_train, X_test, y_train, y_test =
train_test_split(X, y, random_state =
42)
```

The data sets are represented in barh as shown by the Fig. 2.

The logistic regression classifier class was used to build the model. Apply the .train() method to train the model and the .accuracy() method to test the model on the testing data. The data tested by the built model are presented in Table II.

Accuracy is defined as the percentage of evaluation from the students in the testing dataset for which the model was

<BarContainer object of 5 artists>

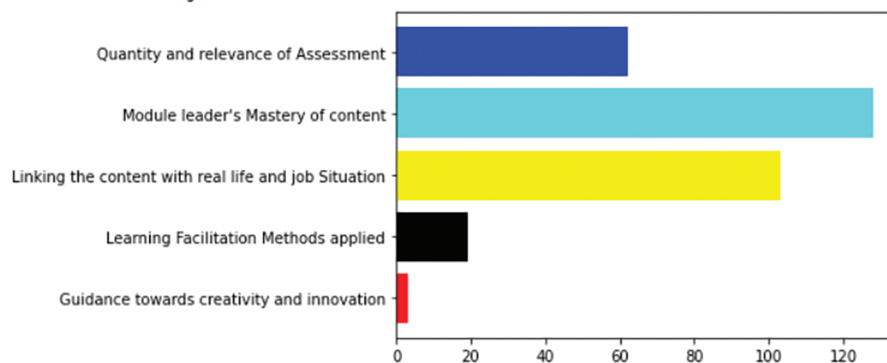


Fig. 2. Visualization of data for training.

TABLE II: REPRESENTATION OF TESTED DATA

	Precision	Recall	f1-score	Support
0	0.68	0.81	0.74	9022
1	0.95	0.9	0.93	36861
Accuracy			0.89	45883
Macro avg.	0.81	0.86	0.83	45883
Weighted avg.	0.9	0.89	0.89	45883

Note: Accuracy: 0.8865157029836759.

correctly able to predict the sentiment. An 88.6% accuracy on the test set is good.

5. CONCLUSION

Module evaluation is important in the life of the academic institution, being implemented is vital in developing quality of module delivered. Machine learning algorithms are key factor in nowadays data analysis and visualization for many institutions not only academic and the results from them are helping in decision making and the future development of the company and institutions.

6. RECOMMENDATIONS

Students Sentiment Analysis using Natural Language Toolkit in Machine Learning for Module Evaluation should be a model and the implementation of this model should be monitored and adjust as necessary.

The study is concentrated on module evaluation using logistic regression; however, a similar study can be carried out in multidimensional area.

CONFLICT OF INTEREST

The authors declare that they do not have any conflict of interest.

REFERENCES

- [1] Cambrian E, Schuller B, Xia YHC. New avenues in opinion mining and sentiment analysis. *IEEE Intell Syst.* 2013;28:15–21. doi: 10.1109/MIS.2013.30.
- [2] Yang L, Li y, Wang JSR. Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon. *IEEE Access.* 2020;8:23522–30. doi: 10.1109/ACCESS.2020.2969854.
- [3] Carosia A, Coelho GSA. Analyzing the Brazilian financial market through Portuguese sentiment analysis in social media. *Appl Artif Intell.* 2019;34(1):1–19. doi: 10.1109/TIA.2022.9667151.
- [4] Capuano N, Greco L, Ritrovato PVM. Sentiment analysis for customer relationship management. *An Increm Appl Intell.* 2020;50(6):1–14. doi: 10.1007/s10489-020-01984-x.
- [5] Sharma S, Daga MGB. Twitter sentiment analysis for brand reputation of smart phone companies in India. In *Proceedings of ICETIT 2019*, pp. 841–52, 2020. doi: 10.1007/978-3-030-30577-2_75.
- [6] Imran A, Daudpotan S, Kastrati ZBR. Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets. *IEEE Access.* 2020;8:181074–90. doi: 10.1109/ACCESS.2020.3027350.
- [7] Chauhan P, Sharma NSG. The emergence of social media data and sentiment analysis in election prediction. *J Ambient Intell Humaniz Comput.* 2020;12:1–27. doi: 10.1007/s12652-020-02423-y.
- [8] Mite-Baidal K, Delgado-Vera C, Solis-Aviles E, Espinoza A, Ortiz-Zambrano JVTE. Sentiment analysis in education domain: a systematic literature review. *International Conference*

on Technologies and Innovation, pp. 285–97, 2018. doi: 10.1080/10494820.2020.1826985.