

Emotion Detection for Software Application Reviews Using Machine Learning

1 st Arvemell Abad <i>Bachelor in science of Computer Engineering, Jose Rizal University</i> arvenell.abad@my.jru.edu	2 nd Patrick Bode <i>Bachelor in science of Computer Engineering, Jose Rizal University</i> Patrick.Bode@my.jru.edu	3 rd James Frederic Dulo <i>Bachelor in science of Computer Engineering, Jose Rizal University</i> jamesfrederic.dulo@my.jru.edu	4 th Benjo Dela Cruz <i>Bachelor in science of Computer Engineering, Jose Rizal University</i> benjo.delacruz@my.jru.edu
--	--	---	---

Abstract - The researchers create a program to determine and detect emotions from software application reviews. Software application reviews tend to appear positive and negative only but with this program software developers are able to know emotions behind the text whether the users are sad, happy, angry or confused with the service they offer.

I. INTRODUCTION

Emotion has an important role in conversations. Because it provides context to the conversation. In the internet era text based conversations are very common from comments, reviews, articles, emails to chat so discerning the emotion based from the text are needed for moderation, analysis and advertisement. An emotion is a subjective state of being that we often describe as our feelings. Emotions result from the combination of subjective experience, expression, cognitive appraisal, and physiological responses (Levenson, Carstensen, Friesen, & Ekman, 1991). (Levenson, Carstensen, Friesen, & Ekman, 1991). However, as discussed later in the chapter, the exact order in which the components occur is not clear, and some parts may happen at the same time.

Recognition of human emotions is a vital phase, which is involved in several applications such as augmented and virtual reality, advanced driver assistance systems, computer interaction, and security systems. People can express their emotions in a variety of ways, including through speech, language, and facial expressions. Furthermore, emotions can be detected based on gaze direction and biosignals including electroencephalogram (EEG) and electrocardiogram (ECG). Human-robot interaction that is intelligent uses emotional expressions (HRI). Therefore, the pupils can study better utilizing this strategy. In human connection, 7% of the affective information is conveyed by words, 38% is conveyed by speech tone, and 55% is conveyed by facial expressions. Therefore, face

emotion analysis can be a solid way to recognize human emotions for HRI applications.

“This study aims to develop an algorithm that could categorize emotion behind sentences.”

Keywords— *text classification, natural language processing, bag of words.*

II. RELATED WORKS

This part presents the perspective and ideas given by different analysts and experts in their separate fields. Likewise, thoughts remembered for this part had helped the researchers in acquainting data that is applicable to the exploration study. Also in this research, a search for Related literature and studies is a huge benefit since it gives a superior comprehension of the examination and the points that are pertinent to the area of study.

According to (Ramalingam et al., 2018) Concepts and general algorithms developed for subjectivity and sentiment analysis are used or modified in emotion detection methods. A lot of different strategies are being tried and tested. However, there are few similarities between many of the approaches.

In the study of (Kao et al., 2009) Finding the relationships between particular input texts and the actual emotions that drive the author to type or write in such styles is all that remains of the emotion detection from text problem. Intuitively, particular surface texts included in the input texts and other deeper inferences are typically required to discover the relationships.

According to (Tao, 2004) Emotions have been a part of human experience for a long time before the development of language, so it is difficult to categorize words according to different states of emotion. There are numerous ambiguities, the majority of which arise in anger or sadness.

In the study of (Herzig et al., 2017) A classification task can be represented as a problem with multiple classes or labels. A one-vs.-rest SVM classifier was used for both types of problems. A classifier then sends the decision function value for each emotion in the training data to a test sample.

According to the study of (Majeed et al., 2020) Emotion can be difficult to extract from text because of the nature of the data. If the words that describe a particular emotion are made explicit in the text, it is easier to identify emotions from the text. Emotion, on the other hand, is usually expressed subtly. There may be multiple emotions in a single piece of writing.

According to (Kaur & R. Saini, 2014) Decision trees are more frequently used to classify literary works, particularly poetry. In the case of poetry, the statistical machine learning approach known as

III. METHODOLOGY

A. Data Collection Techniques

[illegible]

Figure 1 shows a dataset collected and presented using google sheets.

The researchers used neat text to preprocess the datasets. It is used to remove emoji, stop words and irrelevant words on the application reviews. In addition, neat text is used to remove noise in the datasets.

[illegible]

Figure 2 shows codes used in preprocessing using neattext function.

C. Model Classification

1.1 Algorithm

Emotion Detection for Software Application Reviews Using Machine Learning uses logistic regression in predicting or detecting emotions like angry, sad, happy, and confused. It is a natural language processing (NLP) project that can classify text based on the categories stated.

1.2 Flowchart

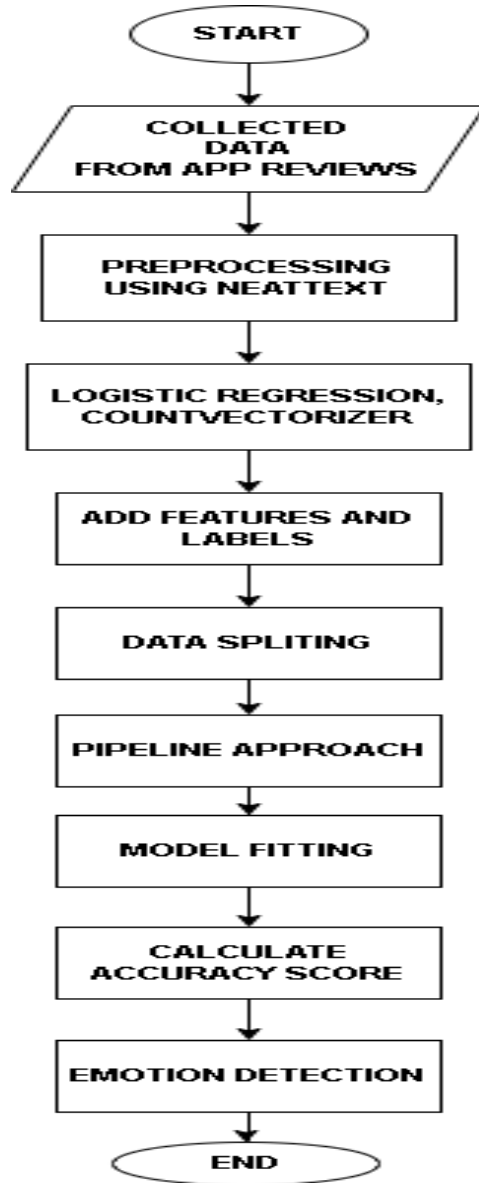


Figure 3: System Flowchart

Figure 3 shows flowchart on how the model is created. It started with collecting data and then loading it to the Python notebook. Next is to clean the

datasets using neatttext as a preprocessing tool and then import important machine learning packages like logistic regression and countvectorizer. Next, add model features and labels so that the data can be split and undergo a pipeline approach. After splitting and fitting the model, calculate accuracy score and then if the model has a good accuracy score then predicting an emotion is the last step.

D. Some Common Mistakes

Common mistakes are collecting datasets. Most comments or application reviews are double sided which means they are giving both positive and negative feedback this will lead to detecting multiple bags of words in one comment and may cause confusion with the program.

IV. EXPERIMENTAL RESULTS

A. Figures and Tables



Figure 4: Angry Prediction

Figure 4 shows text being inputted and then the program detects an angry emotion as an output.



Figure 5: Sad Prediction

Figure 5 shows text being inputted and then the program detects an sad emotion as an output.



Figure 6: Confuse Prediction

Figure 5 shows text being inputted and then the program detects a confused emotion as an output.



Figure 7: Happy Prediction

Figure 7 shows text being inputted and then the program detects a happy emotion as an output.

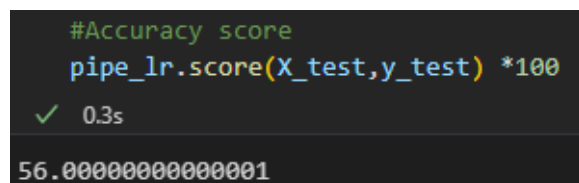


Figure 8: Accuracy Score

Figure 8 shows the accuracy score of the researcher's model and it shows 56.001% accuracy.

V. CONCLUSIONS AND FUTURE WORKS

This research was able to develop a working program that predicts the emotion of the user that depends on the phrase/sentence given by the user. Based on the testing results the program was able to predict the user's emotion through the given sentence, although the data wasn't enough to accurately predict the emotion of the user. We were able to get a 56.01% prediction accuracy using Logistic Regression.

For future works the researchers suggested only getting shorter comments from different sites or apps. Also having a large amount of data will be a huge help in getting a better prediction accuracy. Also, train and fit the model properly.

References

- [1] Graterol, W., Diaz-Amado, J., Cardinale, Y., Dongo, I., Lopes-Silva, E., & Santos-Libarino, C. (2021). Emotion detection for social robots based on nlp transformers and an emotion ontology. *Sensors (Switzerland)*, 21(4), 1–19. <https://doi.org/10.3390/s21041322>
- [2] Haryadi, D., & Putra Kusuma, G. (2019). Emotion Detection in Text using Nested Long Short-Term Memory. In *IJACSA International Journal of Advanced Computer Science and Applications (Vol. 10, Issue 6)*. www.ijacsa.thesai.org
- [3] Herzig, J., Shmueli-Scheuer, M., & Konopnicki, D. (2017). Emotion detection from text via ensemble classification using word embeddings. *ICTIR 2017 - Proceedings of the 2017 ACM SIGIR International Conference on the Theory of Information Retrieval*, 269–272. <https://doi.org/10.1145/3121050.3121093>
- [4] Jain, U., & Sandhu, A. (2015). International Journal of Current Engineering and Technology A Review on the Emotion Detection from Text using Machine Learning Techniques. In 2645| *International Journal of Current Engineering and Technology (Vol. 5, Issue 4)*. <http://inpressco.com/category/ijcet>
- [5] Kao, E. C. C., Liu, C. C., Yang, T. H., Hsieh, C. T., & Soo, V. W. (2009). Towards text-based emotion detection: A survey and possible improvements. *Proceedings - 2009 International Conference on Information Management and Engineering, ICIME 2009*, 70–74. <https://doi.org/10.1109/ICIME.2009.113>
- [6] Kaur, J., & R. Saini, J. (2014). Emotion Detection and Sentiment Analysis in Text Corpus: A Differential Study with Informal and Formal Writing Styles. *International Journal of Computer Applications*, 101(9), 1–9. <https://doi.org/10.5120/17712-8078>
- [7] Majeed, A., Mujtaba, H., & Beg, M. O. (2020). Emotion detection in roman urdu text using machine learning. *Proceedings - 2020 35th IEEE/ACM International Conference on Automated Software Engineering Workshops, ASEW 2020*, 125–130. <https://doi.org/10.1145/3417113.3423375>
- [8] Polignano, M., de Gemmis, M., Basile, P., & Semeraro, G. (2019). A comparison of Word-Embeddings in Emotion Detection from Text using BiLSTM, CNN and Self-Attention. *ACM UMAP 2019 Adjunct - Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, 63–68. <https://doi.org/10.1145/3314183.3324983>
- [9] Ramalingam, V. v., Pandian, A., Jaiswal, A., & Bhatia, N. (2018). Emotion detection from text. *Journal of Physics: Conference Series*, 1000(1). <https://doi.org/10.1088/1742-6596/1000/1/012027>

- [10] Tao, J. (2004). Context based emotion detection from text input. 8th International Conference on Spoken Language Processing, ICSLP 2004, 1337–1340.
<https://doi.org/10.21437/interspeech.2004-329>
- [11] Udochukwu, O., & He, Y. (n.d.). A Rule-Based Approach to Implicit Emotion Detection in Text.