Two Step Approach for Emotion Detection on Twitter Data

Abstract

“Emotional states of individuals, also known as moods, are central to the expression of thoughts, ideas and opinions, and in turn impact attitudes and behavior”. In this paper we have proposed a method which detects the emotion or mood of the tweet and classify the twitter message under appropriate emotional category. Our approach is a two-step approach, it is so called as it uses two approaches for the classification process, one is Rule Based approach and the other is Machine Learning approach. The first approach is the Rule Based Approach (RBA), our minor contributions in this approach are pre-processing, tagging, feature selection and Knowledge base creation. Feature selection is based on tags. Our second approach is Machine Learning Approach (MLA), in this the classifier is based on supervised machine learning algorithm called Naïve Bayes which requires labeled data. Naïve Bayes is used to detect and classify the emotion of a tweet. The output of RBA is given to MLA as input because MLA requires labeled data which we have already created through RBA. We have compared the accuracies of both the approaches, observed that, with the rule based approach we are able to classify the tweets with accuracy around 85% and with the machine learning approach the
Two Step Approach for Emotion Detection on Twitter Data

accuracy is around 88%. Machine learning approach performance is better than rule based approach, the performance has been improved as we have removed the error data while training the model. The approaches are involved with the concepts of Natural Language Processing, Artificial Intelligence, and Machine Learning for the development of the system. Our major contributions in this paper are detection of emotion for non hashtagged data and the labeled data creation for machine learning approach without manual creation.

References

Two Step Approach for Emotion Detection on Twitter Data

multivariable methods. Duxbury Pr.
17. http://sentiwordnet.isti.cnr.it

Index Terms

Computer Science Information Sciences

Keywords

Emotion, Natural Language Processing, POS Tagging