

Hybrid semantic model based on machine learning for sentiment classification of consumer reviews

Palaniraj Rajidurai Parvathy¹, Nagarajan Mohankumar², Rajendran Shobiga³,
Gour Sundar Mitra Thakur⁴, Mamatha Bandaru⁵, Velusamy Sujatha⁶, Shanmugam Sujatha⁷

¹Project Manager, Mphasis Corporation, Chandler, United States of America

²Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India

³Department of Electronics and Communication Engineering, J. J. College of Engineering and Technology, Tiruchirappalli, India

⁴Department of Computer Science and Engineering (AIML), Dr. B. C. Roy Engineering College, Durgapur, India

⁵Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India

⁶Department of Electronics and Communication Engineering, Shree Sathyam College of Engineering and Technology, Sangagiri, India

⁷Department of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India

Article Info

Article history:

Received Feb 27, 2024

Revised Dec 20, 2024

Accepted Jan 27, 2025

Keywords:

Consumer reviews

Ensemble categorization

Lexicon model

Sentiment analysis

Sentiment score

ABSTRACT

Digital information is regularly produced from a variety of sources, including social media and customer service reviews. For the purpose of increasing customer happiness, this written data must be processed to extract user comments. Consumers typically share comments and thoughts about consumable items, technological goods, and services supplied for payment in the modern period of consumerism with simple access to social networking globe. Each object has a plethora of remarks or thoughts that demand special attention due to their sentimental worth, especially in the written portions. The goal of the current project is to do sentiment prediction on the Amazon Electronics, Kindle, and Gift Card datasets. In order to predict sentiment and evaluate utilizing many executions evaluates admitting accuracy, recall, and F1-score, a hybrid soft voting ensemble method that combines lexical and ensemble methodologies is proposed in this study. In addition to calculating a subjectivity score and sentiment score, this study also suggests a non-interpretive sentiment class label that may be used to assess the sign of the evaluations applying suggested method for sentiment categorization. The effectiveness of our suggested ensemble model is examined using datasets from Amazon customer product reviews, and we found an improvement of 2-5% in accuracy compared to the current state-of-the-art ensemble method.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Nagarajan Mohankumar

Symbiosis Institute of Technology, Symbiosis International (Deemed University)

Nagpur Campus, Pune, India

Email: nmkprofessor@gmail.com

1. INTRODUCTION

Numerous sectors for data analysis have emerged as a result of the enormous amount of details created by popular social media platforms. The goal of sentiment analysis (SA), in particular, is to glean user thoughts about a good or service from the relevant text. The raw text must be analyzed using machine learning (ML) algorithms along with natural language processing (NLP) techniques due to the enormous amount of information generated [1]. Sentiment examination made possible through the detail that social media is a necessary device for people and that they regularly distribute their notions. Expressions of