

Submission Summary

Conference Name

Pacific-Asia Conference on Knowledge Discovery and Data Mining

Track Name

Research Track

Paper ID

3

Paper Title

The Amazon Review Spectrum: A Sentiment Analysis Exploration

Abstract

With time, most tasks have become digitized. The growth in the use of the internet along with smart devices brings all decision-making processes to the fingertips. Online shopping has become an easier alternative to in-person shopping now more than ever before. When customers are not able to rely on physical features to confirm a purchase, the next source of reliability comes from online product reviews. Such reviews not only help consumers but also help businesses make predictions about future trends. With the use of NLP, comments in bulk can be categorized with ease. This paper is based on an Amazon product review dataset focusing on gadgets. After tokenizing, in order to prepare the data for the machine learning model, reviews were transformed into numerical values. Ratings, content, and support votes are sectioned using a binary classifier to separate positive and negative reviews. The WordNet Lemmatizer from the NLTK library was used to make semantic associations between words. The NLTK POS tagger was used for nouns, verbs, and adjectives while treebank tags were mapped to WordNet parts of speech. The K-means clustering method was used to combine various reviews with similar sentiments. The Support Vector Machine (SVM) classifier demonstrated the highest pre-optimization accuracy (93.94%) among the various models tested. The SVM's performance increased after some fine-tuning, and it now has an accuracy of 94.08%. Comparatively, the Multinomial Naive Bayes model achieved 93.45% accuracy, while the logistic regression model showed a 93.70% accuracy.

Created on

8/31/2023, 6:59:04 PM

Last Modified

8/31/2023, 6:59:04 PM

Authors

Sadiul Arefin Rafi (Brac University) < sadiul.arefin.rafi@g.bracu.ac.bd > ✓

Annajiat Alim Rasel (BRAC University) < annajiat@gmail.com > ✓

Primary Subject Area

Applications -> Social network analysis

Secondary Subject Areas

Methods and algorithms -> Model selection and evaluation

Submission Files

The Amazon Review Spectrum A Sentiment Analysis Exploration (1).pdf (202.5 Kb, 8/31/2023, 6:58:03 PM)

Submission Questions Response

1. Eligibility for student paper award

Is (Are) the primary author(s) of this work a registered student (students) at the time of submission?

Yes

2. Submission Policy

I understand and confirm that:

1) This submission represents new and original work.

- 2) Concurrent submissions are not allowed.
- 3) This paper (or a substantially similar one) was not published in or accepted to any peer-reviewed journal or conference or workshop with published proceedings, is not currently under review, and will not be submitted to other meetings or publications while under review of PAKDD 2024.
- 4) If the submission is available in arXiv, the last submission date is on or before the date of one month before the submission deadline.
- 5) PAKDD has a strict policy against plagiarism and self-plagiarism. All previously published work must be appropriately cited.
- 6) If this paper is accepted, at least one author will complete the regular registration and attend the conference to present the paper. For no-show authors, the paper will not be included in the proceedings.
- 7) The submission adheres to the format guidelines specified in the call for papers.
- 8) Conflicts of interest are disclosed completely and correctly in the CMT system.

Agreement accepted
