

## ABSTRACT

The development of online stores in Indonesia is going in congruence with the growth of internet users. Customers' trust is the a crucial aspect in online transaction. One of the weaknesses of online stores is the limitation of physical contact between consumers, sellers, and goods. This can affect the level of consumer's confidence. Ratings and reviews are provided to overcome such limitations. In some cases, there is an inconsistency between rating and review, where the polarity between rating and review is contradictory. Handling of the inconsistencies has not been a major focus in the management of rating and review in Indonesian language. This study aims to handle the inconsistency between rating and review in Indonesian language.

The data of rating and reviews are taken from online stores in one of the Indonesian marketplaces. There are two main steps taken to build an inconsistency detection model. The first stage is sentiment analysis. Sentiment analysis is done to handle review data. Sentiment analysis using machine learning approach. The second stage is the classification of ratings and reviews into categories that are consistent or inconsistent. In the second stage, classification using machine learning and classification using rule based are compared.

Based on the evaluation of the research that has been carried out, the proposed approach can detect an inconsistency of rating and review in Indonesian language. The resulting accuracy value is 85.67%.

**Keywords:** rating, review, inconsistency, online store, machine learning, rule based

## INTISARI

Perkembangan toko daring di Indonesia, selaras dengan perkembangan pengguna internet. Kepercayaan menjadi faktor yang penting dalam transaksi daring. Salah satu kelemahan toko daring yaitu keterbatasan kontak fisik antara konsumen, penjual, dan barang. Hal ini dapat mempengaruhi tingkat kepercayaan konsumen. *Rating* dan *review* disediakan untuk mengatasi keterbatasan tersebut. Dalam beberapa kasus, ditemukan inkonsistensi antara *rating* dan *review*, dimana polaritas antara *rating* dan *review* bertolak belakang. Penanganan terhadap inkonsistensi belum menjadi fokus utama dalam pengelolaan *rating* dan *review* Bahasa Indonesia. Penelitian ini bertujuan untuk melakukan penanganan terhadap inkonsistensi antara *rating* dan *review* Bahasa Indonesia.

Data *rating* dan *review* diambil dari toko daring yang ada di salah satu *marketplace* Indonesia. Terdapat dua tahap utama yang dilakukan untuk membangun model deteksi inkonsistensi. Tahap pertama yaitu analisis sentimen. Analisis sentimen dilakukan untuk menangani data *review*. Analisis sentimen dilakukan dengan menggunakan pendekatan *machine learning*. Tahap kedua yaitu klasifikasi *rating* dan *review* ke dalam kategori konsisten atau inkonsisten. Pada tahap kedua dilakukan perbandingan antara klasifikasi menggunakan *machine learning* dan klasifikasi menggunakan *rule based*.

Berdasarkan hasil evaluasi penelitian yang telah dilakukan, pendekatan yang diajukan dapat mendeteksi adanya inkonsistensi *rating* dan *review* Bahasa Indonesia. Nilai akurasi yang dihasilkan sebesar 85,67%.

**Kata kunci** -- *rating*, *review*, inkonsistensi, toko daring, *machine learning*, *rule based*