A Feature-Enriched Machine Learning Approach to Deceptive Review Detection in Indian E-Commerce

Janhvi Pralhad Babar[1], Shruti Manoj Gotral[2],
Pranita Dnyaneshwar Gaikwad[3], Prof. N. S. Kharatmal[4]
Student, Computer Science and Engineering [1,2,3]
Lecturer, Computer Science and Engineering[4]
Matsyodari Shikshan Sanstha College of Engineering and polytechnic, Jalna, India
babar.janhvi333@gmail.com[1] , shrutigotral59@gmail.com[2] , pranitagaikwad2007@gmail.com[3] ,
nanditakhartmal27@gmail.com[4]

Abstract: *With India's e-commerce sector hitting record growth, the "star rating" has become a vital digital currency for shoppers on Flipkart and Amazon India. Yet, this total reliance on community feedback has triggered a new, highly sophisticated wave of review fraud. Modern fake reviews have evolved past obvious bot templates; they now replicate the specific tone, Hinglish vocabulary, and cultural nuances of genuine Indian buyers so accurately that traditional detection tools have become obsolete.*

This research presents the Feature-Enriched Machine Learning (FEML) framework, designed specifically for the complexities of the Indian market. Moving away from one-dimensional analysis, our model "interrogates" reviews through a three-layer process: (1) Syntactic/Semantic Cues for linguistic patterns, (2) Behavioral Metadata to flag anomalies like post-frequency spikes, and (3) Sentiment Consistency to catch "rating-text" mismatches. Testing against diverse, high-stakes Indian product datasets using a Random Forest and Gradient Boosting ensemble, the framework achieved a 94.2% detection accuracy. Our results prove that text-only analysis is no longer enough; unmasking deceptive intent now requires a deep dive into the reviewer's long-term digital footprint..

Keywords: Deceptive Review Detection, Machine Learning, Indian E-Commerce, Opinion Spam, Natural Language Processing, Feature Engineering