

# Building Sentiment Analysis Model and Compute Reputation Scores in E-Commerce Environment Using Machine Learning Techniques

Elshrif Ibrahim Elmurngi, École de Technologie Supérieure, Montreal, Canada

Abdelouahed Gherbi, École de Technologie Supérieure, Montreal, Canada

## ABSTRACT

Online reputation systems are a novel and active part of e-commerce environments such as eBay, Amazon, etc. These corporations use reputation reporting systems for trust evaluation by measuring the overall feedback ratings given by buyers, which enables them to compute the reputation score of their products. Such evaluation and computation processes are closely related to sentiment analysis and opinion mining. These techniques incorporate new features into traditional tasks, like polarity detection for positive or negative reviews. The “all excellent reputation” problem is common in the e-commerce domain. Another problem is that sellers can write unfair reviews to endorse or reject any targeted product since a higher reputation leads to higher profits. Therefore, the purpose of the present work is to use a statistical technique for excluding unfair ratings and to illustrate its effectiveness through simulations. Also, the authors have calculated reputation scores from users’ feedback based on a sentiment analysis model (SAM). Experimental results demonstrate the effectiveness of the approach.

## KEYWORDS

E-Commerce (EC), Logistic Regression (LR), Reputation Systems, Sentiment Analysis Model (SAM)

## 1. INTRODUCTION

E-commerce has become one of the major way of shopping for products ranging from simple electronics to valuable items. Almeroth, and Zhao (2010) where collective unfair ratings are referred to as collusion Sun and Liu (2012); Swamynathan et al. (2010) and are more complicated and much difficult to detect than the single unfair ratings Sun and Liu (2012). For that reason, in the present work, we are focusing on understanding and identifying unfair rating scores, all good reputation problems, and collusion and manipulation detection. In online shopping, the customers often depend on other customers’ feedback posted through a rating system, before deciding on buying a product (Mukherjee, Liu, & Glance, 2012). Online feedback-based rating systems, also known as online reputation systems, are systems in which users provide ratings to items they bought. Based on the feedback of the product, the consumer decides whether to buy the product or not. This motivates the seller to promote or demote a product of their interest depending on their competitors’ product, by posting rating scores which are unfair (Brown and Morgan, 2006; Harmon, 2004). For example, fraudulent sellers may try to increase their income by submitting positive feedback which increases their product rating (Harmon, 2004). In addition, occasional sellers on eBay boost their reputations

DOI: 10.4018/IJOCI.2020010103

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

unfairly by selling or buying feedbacks (Brown and Morgan, 2006). Unfair rating scores are given singularly or collectively Swamynathan,

In the present work, we are focusing on understanding and identifying unfair rating scores, all good reputation problems, and collusion and manipulation detection.

An online rating system needs truthful feedback in order to work properly. An important part of a rating system is creating honest and representative feedback. It should not only have qualitative, quantitative facts opinion-based process, but also should detect situations where some users may try to mislead the system by providing unfair positive and negative reviews, which is likely to lead to collusion and manipulation. Figure 1 Generic process of a reputation system and consists of three main components: feedback collection, computation and rating scores presentation (Noorian and Ulieru, 2010; Sun and Liu, 2012). Feedback Collection is responsible for collecting feedback from community and feeding it to the rating system. A feedback is the opinion of an evaluator on the quality of an item or a person. Generally, a feedback can be expressed either as a number, such as a textual review, or a rating score (Adler, de Alfaro, Kulshreshtha, and Pye, 2011). A numeric feedback can be a negative or positive number chosen from either a discrete list of options, for example, an integer choice between 1 and 5 representing quality of an item. Feedback computation is responsible for computation of all feedbacks received from evaluators to calculate correct rating scores for people and items. Investigating the feedback computation part after using sentiment classification algorithm is the focus of this study.

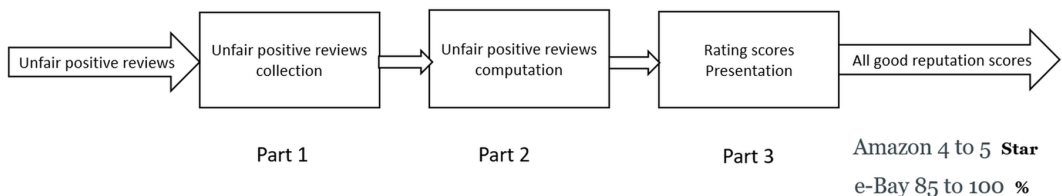
According to Reyes-Menendez, Saura, and Martinez-Navalon (2019), it is improbable for all consumers to be satisfied with a product, very positive ratings might be interpreted by users as incredible information. Consequently, it is recommended that companies ensure that the feedback made to them on platforms such as TripAdvisor, contains only actual comments, as well as several less positive comments; this can generate greater customer credibility and reliability of an offered product.

Figure 2 shows how unfair positive and negative reviews affects negatively on reputation scores computation. If the collection part collects unfair positive and negative reviews, the computation part will compute unfair positive and negative reviews and then the rating scores presentation part will present unfair rating scores as output.

Both the research community and the e-commerce industries has accepted unfair reviews to be a crucial challenge to the e-commerce industry Feng, Xing, Gogar, and Choi (2012) and Breure (2013); Sussin and Thompson (2012). Any (positive or negative) review that is an unfair review and not an actual consumer's honest opinion will affect reputation scores negatively.

The main objective of this study is to offer a novel and comprehensive solution for designing a new model to obtain the most accurate reputation system, which addresses the existing issues, such as collusion and manipulation and the "all good reputation" issue that is being currently encountered by reputation systems. While applied reputation models currently rely mainly on the overall ratings of items, they do not involve customer reviews in their assessment. Conversely, few of the reputation models focus only on the overall reviews of products without considering the ratings provided by the customers. This research aims to compute feedback rating based on feedback reviews. Subsequently, in order to get accurate reputation scores, we propose a simple calculation method that calculates

Figure 1. Generic process of "all good reputation" problem



29 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/article/building-sentiment-analysis-model-and-compute-reputation-scores-in-e-commerce-environment-using-machine-learning-techniques/243677](http://www.igi-global.com/article/building-sentiment-analysis-model-and-compute-reputation-scores-in-e-commerce-environment-using-machine-learning-techniques/243677)

## Related Content

---

### Applications in Multiobjective, Constrained and Minimax Problems

E. Parsopoulos Konstantinos and N. Vrahatis Michael (2010). *Particle Swarm Optimization and Intelligence: Advances and Applications* (pp. 245-268).  
[www.irma-international.org/chapter/applications-multiobjective-constrained-minimax-problems/40638](http://www.irma-international.org/chapter/applications-multiobjective-constrained-minimax-problems/40638)

### Creative Accelerated Problem Solving (CAPS) for Advancing Business Performance

Cookie M. Govender (2021). *Handbook of Research on Using Global Collective Intelligence and Creativity to Solve Wicked Problems* (pp. 84-109).  
[www.irma-international.org/chapter/creative-accelerated-problem-solving-caps-for-advancing-business-performance/266781](http://www.irma-international.org/chapter/creative-accelerated-problem-solving-caps-for-advancing-business-performance/266781)

### Drilling Optimization via Particle Swarm Optimization

T. O. Ting and T. S. Lee (2012). *International Journal of Swarm Intelligence Research* (pp. 43-54).  
[www.irma-international.org/article/drilling-optimization-via-particle-swarm/67537](http://www.irma-international.org/article/drilling-optimization-via-particle-swarm/67537)

### The Use of Evolutionary Algorithm-Based Methods in EEG Based BCI Systems

Adham Atyabi, Martin Luerssen, Sean P. Fitzgibbon and David M W Powers (2013). *Swarm Intelligence for Electric and Electronic Engineering* (pp. 326-344).  
[www.irma-international.org/chapter/use-evolutionary-algorithm-based-methods/72835](http://www.irma-international.org/chapter/use-evolutionary-algorithm-based-methods/72835)

### Ethics and Natural Resources Management: On Commons, Anti-Commons, Illegal Behavior, and Bureaucracy

Manuel Pacheco Coelho, José António Filipe and Manuel Alberto M. Ferreira (2012). *International Journal of Organizational and Collective Intelligence* (pp. 56-71).  
[www.irma-international.org/article/ethics-and-natural-resources-management/100003](http://www.irma-international.org/article/ethics-and-natural-resources-management/100003)