


# Crowdsourced Social Media Reaction Analysis for Recommendation

Jaiprakash Vinodkumar Verma, Institute of Technology, Nirma University, India

 <https://orcid.org/0000-0001-6116-1383>

Sudeep Tanwar, Institute of Technology, Nirma University, India

Sanjay Garg, Institute of Technology, Nirma University, India

Abhay Dinesh Rathod, Institute of Technology, Nirma University, India

## ABSTRACT

A pre-analysis is always important for crucial decision making in many events where reviews, feedback, and comments posted by different stakeholders play an important role. Summaries generated by humans are mostly based on abstractive summarization. It sometimes changes the meaning of the text. This paper proposes a customized extractive summarization approach to generate a summary of large text extracted from social media viz. Twitter, YouTube review, feedback, comments, etc. for a movie. The proposed approach where PageRank with k-means clustering was used to select representative sentences from a large number of reviews and feedback. Cluster heads were selected based on the customization of PageRank. The proposed approach shows improved results over the graph-based TextRank approach with and without synonyms. It can be applied to predict trends for items other than movies through the social media platform.

## KEYWORDS

Co-Selection, Extractive Text Summarization, Machine Learning, Opinion Mining, PageRank, Recommendation System, TextRank

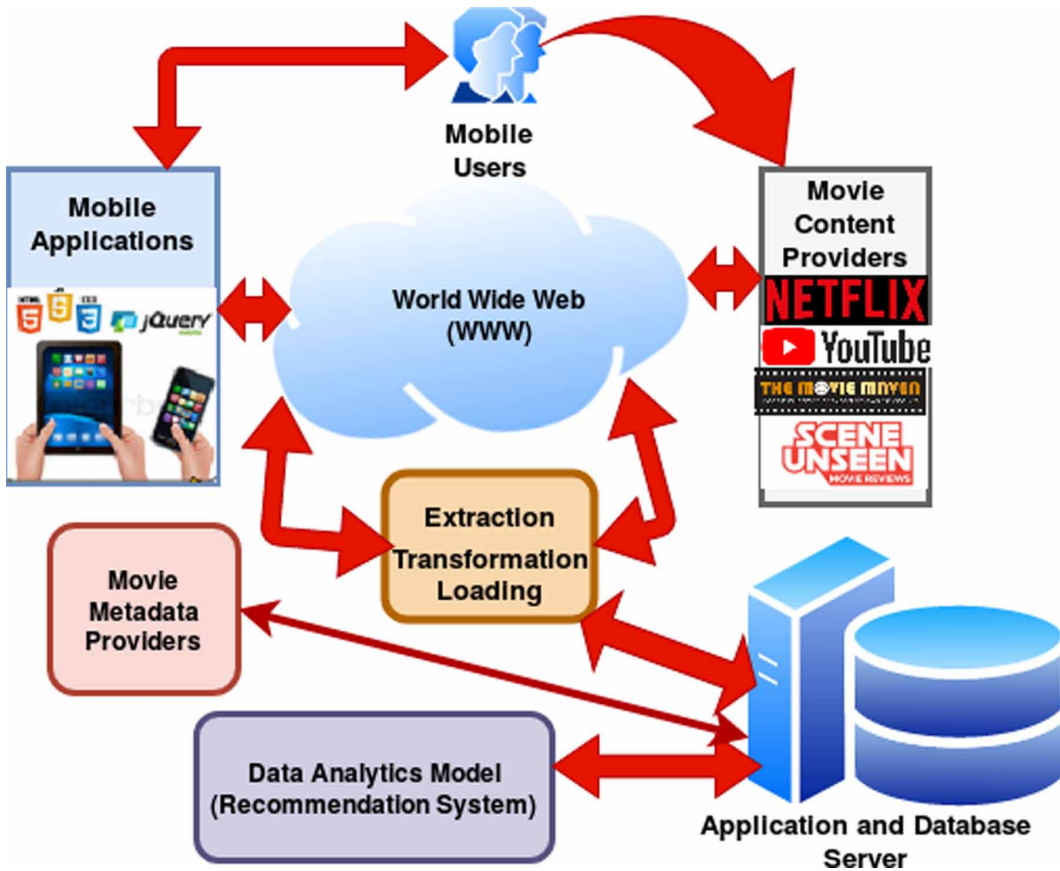
## 1. INTRODUCTION

Text summarization plays a significant role in information retrieval and data processing systems. According to Forbes, an average of 2.5 Quintillion bytes of data are produced every day and 90% of these data growth has been seen in the last two years (Dalwadi, 2017). It becomes necessary to store this data in a structured format and analyze for business needs. The methods like PageRank of information retrieval based on graph data structure can be used for text summarization (Shams B., 2018). It is always a challenge for the data science community to analyze, evaluate and find actionable insights from a large dataset. Handling such huge data and extracting the meaning from it in a short time such that it gives maximum crust is a challenging job. The extractive text summarization can be used to analyze unstructured data extracted from social media sites, reviews and feedback systems. It applies methods like stemming, lemmatization, and feature extraction through Tokenization and Vectorization (Singh J., 2018). Machine learning techniques based on text summarization methods are discussed in (Allahyari M., Dalwadi, B., 2017).

A pre-analysis is always crucial for the decisions on a release date, advertising strategy, promotion techniques. As per (Singh J., 2018) promotional strategy of movies is based on how the public is

DOI: 10.4018/IJKSS.2021010101

Figure 1. Architecture of recommender systems



reacting on trailer launches. Twitter also provides a universal platform to everyone for sharing their thoughts in the form of tweets. Twitter handlers provide tweets related to a particular keyword of a movie like a title, actor, director, etc. As per (Fang, C., 2017), each comment and tweet can be considered as a sentence. and various sentence scoring algorithms applied for text summarization to identify representative sentences. Here, an algorithm such as CoRank and CoRank+ can be used for word-sentence scoring techniques. Also, there are several algorithms present and it has been seen that they show different behavior on the various dataset (Ferreira, R., 2013, Gupta, D. P. N., 2012, Singh, D. P., 2013). Movie RS shows in Figure 1 extract data from different websites like Netflix, YouTube, scene unseen, etc. These texts are posted by different users through various web and mobile applications. Click streams for watching different movie trailers and advertisements also become a crucial data source for RS (Xu, Y., Zhang, F., 2019).

## 1.1. Background

Many online movie streaming sites usages, Recommendation System (RS) to recommend movies to their potential users. As per (Wei, S., Zheng, X., 2016) in RS the prime problem is to deal with a cold start since not all the users are ready to review and rate the content they watch. Da Costa *et al.* (da Costa, A. F., 2019) presents RS like EcoRec, a hybrid approach of two or more algorithms that takes a single input dataset, combines the output of both into a single ensemble. RS analyses the emotive response and social media trends of the movie in the crowd to make major decisions for

17 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/crowdsourced-social-media-reaction-analysis-for-recommendation/271419](http://www.igi-global.com/article/crowdsourced-social-media-reaction-analysis-for-recommendation/271419)

## Related Content

---

### Meta-Synthesis Knowledge System: Basics and Practice

Gu Jifa (2012). *Systems Approaches to Knowledge Management, Transfer, and Resource Development* (pp. 132-146).

[www.irma-international.org/chapter/meta-synthesis-knowledge-system/68215](http://www.irma-international.org/chapter/meta-synthesis-knowledge-system/68215)

### Research on the Cluster-Level Architectural Knowledge: Exploration of Six Content Dimensions of Architectural Knowledge based on a Case Study

Hao Li, Hong-jie Zhang and Kai-jing Zhou (2012). *International Journal of Knowledge and Systems Science* (pp. 53-66).

[www.irma-international.org/article/research-cluster-level-architectural-knowledge/75332](http://www.irma-international.org/article/research-cluster-level-architectural-knowledge/75332)

### Autonomic QoS Optimization of Real-Time Internet Audio Using Loss Prediction and Stochastic Control

Lopamudra Roychoudhuri and Ehab S. Al-Shaer (2012). *Technological Innovations in Adaptive and Dependable Systems: Advancing Models and Concepts* (pp. 239-260).

[www.irma-international.org/chapter/autonomic-qos-optimization-real-time/63585](http://www.irma-international.org/chapter/autonomic-qos-optimization-real-time/63585)

### Solving Multi-Objective Multicast Routing Problem Using a New Hybrid Approach

Mohammed Mahseur, Abdelmadjid Boukra and Yassine Meraihi (2018). *International Journal of Applied Evolutionary Computation* (pp. 22-36).

[www.irma-international.org/article/solving-multi-objective-multicast-routing-problem-using-a-new-hybrid-approach/214894](http://www.irma-international.org/article/solving-multi-objective-multicast-routing-problem-using-a-new-hybrid-approach/214894)

### A Dual-Role Collaborative Learning Support System for Simultaneous Speaking Acquisition in English and Japanese

Anh Bui and Kazushi Nishimoto (2022). *International Journal of Knowledge and Systems Science* (pp. 1-19).

[www.irma-international.org/article/a-dual-role-collaborative-learning-support-system-for-simultaneous-speaking-acquisition-in-english-and-japanese/298709](http://www.irma-international.org/article/a-dual-role-collaborative-learning-support-system-for-simultaneous-speaking-acquisition-in-english-and-japanese/298709)