Chapter 102
Integration of Different Analytical Concepts on Multimedia Contents in Service of Intelligent Knowledge Extraction

Goran Klepac
University College for Applied Computer Engineering Algebra, Croatia & Raiffeisenbank Austria, Croatia

ABSTRACT

This chapter will introduce methodology how to use analytical potential of multimedia contents like YouTube, Bing Videos or Vimeo for discovering behavioral consumer characteristics. Chapter will also show how to consolidate unstructured text data sources from blogs and Twitter with revealed knowledge from multimedia contents for better understanding consumer habits and needs. For this purposes Social Network Analysis will be used as well as text mining techniques on different internet data sources. Presented methodology has practical value where information about customer behavior, preferences and changes in preferences during different time periods is valuable information for campaign planning, campaign management and new product development. Presented methodology also captures different techniques for data crawling from different internet resources, as well as analytical consolidation of revealed results which aim is better understanding of client behavior.

INTRODUCTION

Sites with video content like YouTube, Bing Videos or Vimeo, could be wealth source of information for discovering behavioral characteristics of their consumers. Behavioral characteristics, preferences, profiles, segments, clusters can be discovered by analyzing crawled data from mentioned sites.

Consumer behavior and preferences could be discovered regarding watching preferences.

DOI: 10.4018/978-1-5225-1759-7.ch102

Initial watcher population can be extracted by snowball sampling where root node is some specific user which has published some specific video material, or based on some specific keyword.

Initial population is base for deeper investigation regarding consumer segmentation, consumer behavior recognition and pattern recognition process.

For that purpose, key role has social network analysis (SNA) which helps us find interesting patterns and segments among specific watcher population. SNA gives us opportunity for discovering hidden influencer and leaders among network of watched videos. That leads us on giving answers on questions like: "Which type of videos, consumers like to watch after watching some aimed video? ", "Is there any hidden connections between different video segments regarding consumer preferences?", "Which video/videos within population acts key role for consumer motivation for watching videos from another segment (genre) ? ".

Another valuable source of information is consumer comments associated with video content. That fact leads us on natural language processing area. Recognized patterns and segments via SNA can be additionally analyzed through text mining.

Chapter gives framework and solution for analyzing video consumer behavior using social network analysis and text mining with technical details how to realize proposed solution.

Integral part of the chapter is complete case study from Croatian market. Case study will show proposed framework in action in situation before premiere of film "Fifty shades of gray".

Analysis based on previously explained methodology has been used. Central point of analysis was keyword "*Pedeset nijansi sive*" (Croatian translation of "Fifty shades of gray") on YouTube for data crawling.

As a result, analysis showed interesting patterns and preferences based on consumer behavior, which shows what average consumer interested in forthcoming movie also prefers to watch, and which is focus of his interest beside aimed content.

Same analysis has been performed on YouTube content after mentioned movie started to play in cinema and analysis shows new patterns and changes in consumer behavior.

Text mining analysis has been performed in parallel on text content associated with videos, and those types of analyzed data also shows interesting patterns valuable for market planning.

Same analysis also used Twitter data with the same keyword "*Pedeset nijansi sive*" (Croatian translation of "Fifty shades of gray").

Captured data from Twitter was used for conducting SNA and text mining. At the end, different results were integrated and used for marketing strategy.

Chapter also shows usage of different open source platforms, programming languages, program libraries in service of realization proposed framework and methodology.

Scientific contribution of proposed methodology is in integration of different analytical concepts on multimedia contents in service of intelligent knowledge extraction, which can be used in marketing strategy development.

Background

Understanding of customer behaviour is a key factor of market success, especially in competitive market conditions (Berry, 2000; Giudici, 2003; Giudici, 2009).

Extracting important behavioural information from transactional customer data, and enabling better decision-making throughout an organization is one of the aims when a company wants to understand