

USING SNS DATA IN PREDICTION OF TV VIEWERSHIP

Ook Lee, Department of Information System, Hanyang University, Seoul Korea

dx.doi.org/10.18374/JABE-13-3.8

ABSTRACT

Social Network Sites (SNS), the web-based services, to construct individual profile either publically or privately, share the connections formed among individual relations, and support interactivity brought from those connections (Boyd and Ellison 2008). The state of the usage of SNS by type in Korea is as follows; blog 84.4% , community 74.6% , mini-homepage 68.3% , profile-based service 18.4% , and micro-blog 12.8% in descending order (Korea Communications Commission and Korea Internet Security Agency 2012). Since the analyzing of the massive content data and related network information through SNS became possible, efforts to find and utilize new meaning have recently grown active as well. Thus, as "Big data era" advents, firms are maximizing profit through customized services with big data analysis. Big data imply database with mega-capacity that exceeds the scope of existing database that can store, manage, and analyze. It is the formal or informal data that cannot be collected, stored, retrieved, analyzed, and visualized through existing methods or tools (McKinsey Global Institute 2011). Using big data can provide customized context-based services to each customer and improve customer satisfaction. For example, on December 18th - the day before Korean presidential election, the ratio of people's data search amount on Google for the two candidates were 52.6% for Park, and 47.3% for Moon, which closely matched the actual ratio of 51.6% to 48% (www.chosun.com accessed 2012). The behavior consuming media is evolving from consuming through individualized media to social consumption, meaning participation and sharing. The viewers no longer remain just 'passive viewers'. They also give out opinions while viewing programs and share information or feeling about the program before and after viewing. Such "Social TV" is spreading out, and therefore, complementing relation needs to be formed in addition to existing statistical survey by catching real-time feedbacks through SNS or blog analysis to understand the trend. If influencing power of on-line data as independent variable can be predicted for viewing rate data, then the viewing rate prediction for contents that takes costly production budget and establishing a corresponding marketing strategy will be much easier. Nowadays, it covers approximately 0.2 to 0.5 million dollars per episode to produce a mini-series drama in Korean media industry. One title usually consists of 16 to 20 episodes. Thus, total investment takes about 3 to 9.3 million dollars. So far, the only quantitative measurement tool that broadcasting organizations, producing companies, merchandising providers, and advertisers use to measure the outcome of their investment production is viewer rating data. With "Han-ryu," the Korean culture power spreading globally through K-pop and Korean drama, an analyzing system that can diminish investors' risk, maximize the marketing effect, and swiftly execute strategy into practice will contribute greatly in producing high-quality contents. In Korea, it is pretty rare that television programs are fully produced in advance. Most of the programs go through final editing by episodes about a week before broadcasting. If marketability of contents can be grasped, it will be more convenient to determine (1) whether to stick to the original plan and invest persistently, (2) whether to adjust the investment plan according to viewers' response, or (3) whether to completely put a stop to investment or broadcasting in short amount of time. The purpose of this study is to predict viewers' responses that can be evaluated as viewing rate by analyzing big data buzzword pattern related to the contents created and accumulated on-line during or after the broadcasting of programs. The construction of this study begins with literature review and suppositions setting from the attributes of gathered data. Analyzing tools and the contents that are to be analyzed are presented in the third section. In the fourth section, reliability analysis, primary component analysis, and regression analysis are conducted for the model. Finally, the last section summarizes the results and reaches the conclusion of this study with proper suggestions.

Keywords: