

# Which Scripts Are More Profitable?: A Topic Model Analysis for the Movie's Green-lighting process

**Jongdae Kim**

Seoul National University

**Youseok Lee**

Myongji University

**Inseong Song**

Seoul National University

Cite as:

Kim Jongdae, Lee Youseok, Song Inseong (2020), Which Scripts Are More Profitable?: A Topic Model Analysis for the Movie's Green-lighting process. *Proceedings of the European Marketing Academy*, 49th, (62447)

Paper from the 49th Annual EMAC Conference, Budapest, May 26-29, 2020.



# Which Scripts Are More Profitable?: A Topic Model Analysis for the Movie's Green-lighting process

## **Abstract**

Starting a new movie production is risky. The budget is huge and the profitability of the box office varies largely across movies. Within the same context, although the process to decide which scripts are allowed to move forward to pre-production, known as the green-lighting process, is important for movie studios, this has a lot of difficulties because there are few empirical evidence so that they have to depend on experts' experiences. In this paper, we develop an empirical approach based on text mining to evaluate movie scripts' profitability. First, we use latent Dirichlet allocation to find latent topics from scripts and extract probabilities of scripts assigned to each topic as an explanatory variable. From the regression and the classification based on a movie's return-on-investment, we find the effect of probabilities of specific topics are significant. We expect our model could help studios to make more profitable and systematic decisions in the green-lighting process.

**Keywords:** *New Product Development; Entertainment Industry; Latent Dirichlet Allocation*

**Track:** Innovation Management & New Product Development