Online Educational Video Recommendation System Analysis

Parvathi R.
Vellore Institute of Technology, Chennai, India

Aarushi Siri Agarwal
Vellore Institute of Technology, Chennai, India

Urmila Singh
Vellore Institute of Technology, Chennai, India

INTRODUCTION

Recommendation systems usually filter large amounts of available data and select items that are most likely to be interesting and attractive to users. Recommendation methods are classified into three categories: content-based, collaborative filtering and hybrid methods. The most preferred and popular among the three methods is content-based methods, it looks at the user's history and suggests other items in the same or a similar category to increase user retention and increase in customer base. For example, news recommendations find similarities by considering words or terms in articles.

Content-based filtering is the use of certain features on the basis of likes, comments, reactions and explicit feedback to escalate recommendation for other items. Collaborative filtering uses user ratings and gives more personalized recommendations (Ni et al., 2022). Hybrid method is a combination of content based and collaborative filtering methods.

For generations, humans have looked outwards at others for inspiration, individuals of great intellect and accomplishment have been regarded as heroes and model citizens. However, this kind of admiration has not always been shared equally amongst the diverse set of people in the world. As such, with growing awareness, it is often pondered whether the data we are recommended to solve the issue of finding what we need in the ever growing products and services options online, also known as information overload (Beede et al., 2011) is subjected to any form of bias. A particular focus is on popularity bias caused by recommendation systems (Boring, 2017) leading to extreme disparity. Popularity bias results in the popular videos that have more views and likes to become even more popular while the long-tailed videos remain long-tailed. By investigating user ratings of speakers in TED talks, a popular lecture series that are widely shared for approachable discussions from various experts. The aim is to explore whether a platform primarily served to distribute educational content is subject to potential unconscious biases (Veletsianos et al., 2022) and if certain types of content are systematically more recommended than others. In order to achieve this aim we try finding the different correlations between parameters of uploaded talk videos and visually represent these correlations for better understanding (Pedersen and Duin, 2022).

DOI: 10.4018/978-1-7998-9220-5.ch093

BACKGROUND

Systems designed based on content-based recommendation explicitly analyze parameters, considering feedback, descriptions of previously rated items, and work on further building a model based on the candidate specification and similarity metrics. Several channels can be used up for generating relevant feedback and specifications required. New interesting items are then recommended on the basis of the structure model developed. Highly positive and accurate recommendations require good hand-engineered features. Semantic analysis (lexicons and ontologies) are used by other members of Content-based RSs for enhanced and accurate item representation. Furthermore (Kempe et al., 2003) discussed one of the most popular algorithms for extracting clusters in a graph is proposed by New-man and Girvan (Newman and Girvan, 2004), which is based on modularity. (Zhou et al., 2016), developed a modularity based graph clustering approach, which can help users discover common interests of other users quite effectively. It obtains an undirected weighted tag-graph for each user. One of other methods is the Content-based video retrieval (CBVR) technique, which has been widely used in content-based video lecture recommendation systems (Zhou et al., 2016). OCR, ASR techniques, and folksonomy are very widely used in annotation tasks for information retrieval for user preference. In most of the related works, the solidity and consistency problems are caused by varying accuracy of different analysis engines (Yang and Meinel, 2014), which has not been thoroughly discussed. (Wingrove 2022; Holland et al., 2022).

Artificial intelligence like Iris AI is known for indexing academic resources. The system further recommends scientific papers and speeches using a model trained with TED Talks videos which is related to the field of interest of this research paper. In a paper by Yang and Meinel, 2014), the authors applied automatic video segmentation and key-frame detection for the video lecture content navigation.

Latent Dirichlet Allocation (LDA) (Shiokawa et al., 2013) has also been used to ðnd the topics in a particular website and those in a particular video. Wang and Blei combined probabilistic topic modeling with collaborative ðltering to recommend scientiðc papers. Chen, Cooper, Joshi, and Girod have proposed a Multi-modal Language Model (MLM) that uses latent variable techniques to explore the co-occurrence relation between multi-modal data (Chen et al., 2014).

Clauset et al. (2004) proposed a greedy modularity-based algorithm, called CNM which is one of the most widely used methods recently. Blondel et al. (2008) proposed an efficient greedy algorithm BGLL. To the best of our knowledge, BGLL is representative of the state-of-the-art algorithm; it achieves fast clustering with higher modularity than the other algorithms. In contrast to CNM, it computes the modularity gain only for the adjoined vertex pairs as a local maximization (Newman and Girvan, 2004).

In the literature survey, a variety of content-based recommendation algorithms have been proposed. A traditional example is the "k-nearest neighbours" approach (KNN) that computes the preference of a user for an unknown item by comparing it against all the items known by the user in the catalogue (Tang and Yu, 2022). Similarity of the known to the unknown items, helps in predicting the preference score. Dot Product or similarity metrics is a common parameter to mark the similarity. Relevant works include models for user's interest and recommendation based on Bayesian approach, IR (Information Retrieval), or a product of IR, i.e., Relevance Feedback method.

FEASIBILITY STUDY

The analysis parameters for an online video recommendation system could be of any form: title, main speaker, film date, opening words related to the videos. Here, we specifically study and build our model

Related Content