Movie Genre Preference Prediction Using Machine Learning for Customer-Based Information

Authors : Haifeng Wang, Haili Zhang

Abstract : Most movie recommendation systems have been developed for customers to find items of interest. This work introduces a predictive model usable by small and medium-sized enterprises (SMEs) who are in need of a data-based and analytical approach to stock proper movies for local audiences and retain more customers. We used classification models to extract features from thousands of customers’ demographic, behavioral and social information to predict their movie genre preference. In the implementation, a Gaussian kernel support vector machine (SVM) classification model and a logistic regression model were established to extract features from sample data and their test error-in-sample were compared. Comparison of error-out-sample was also made under different Vapnik–Chervonenkis (VC) dimensions in the machine learning algorithm to find and prevent overfitting. Gaussian kernel SVM prediction model can correctly predict movie genre preferences in 85% of positive cases. The accuracy of the algorithm increased to 93% with a smaller VC dimension and less overfitting. These findings advance our understanding of how to use machine learning approach to predict customers’ preferences with a small data set and design prediction tools for these enterprises.

Keywords : computational social science, movie preference, machine learning, SVM

Conference Title : ICCSIE 2017 : International Conference on Computer Science and Information Engineering

Conference Location : San Diego, United States

Conference Dates : December 18-19, 2017