Prediction of Eudaimonic and Hedonic Movie Characteristics From Subtitles

Elham Motamedi[1], Marko Tkalcic[1]

[1]University of Primorska, Koper 6000, Slovenia

Abstract

Personalization and explaining the recommendations of recommender systems (RSs) have recently gotten more attention from researchers. We would like to use eudaimonic and hedonic perceptions of users from movies to personalize the movie recommendations. To achieve this goal, first, we need to predict the movies' eudaimonic/hedonic quality features. This research study presents the pipeline and preliminary results for predicting movies' eudaimonic and hedonic characteristics.

Keywords

personalization, recommender system, eudaimonic perception, hedonic perception

1. Introduction

Movie recommender systems have been studied extensively in the last two decades [1]. Roughly, we distinguish three groups of algorithmic approaches: (i) content-based (CBR), where the recommendations are based on item characteristics, (ii) collaborative, where the predictions are based on similarities between users and items [2] and (iii) hybrid, where the recommendations are based on item characteristic and similarities between users and items. This work addresses CBR. The proposed idea can also be exploited in hybrid approaches. Related work has exploited a wide variety of item characteristics for movie recommendations, such as genre, year, actors, directors, etc. While these characteristics have been proven to contain information about user preferences, we conjecture that item characteristics that have a more intimate reflection of the user perception of the item might carry more information about user preferences, hence leading to better recommendations. We aim at investigating a novel set of item characteristics, namely the hedonic and eudaimonic qualities of a movie. The eudaimonic quality is related to the meaning or the goal of the life that one pursues, while the hedonic quality is more associated with the plain pleasure that one experiences [3, 4, 5]. The more meaningful a multimedia item is for a user, the higher the value of eudaimonic perception is for that person. Similarly, the hedonic perception score is the degree of plain pleasure that the user experience while consuming the multimedia item. Therefore, each multimedia item and in particular each movie, can get a score of eudaimonic/hedonic perceptions for each user. Each item has a distribution of eudaimonic/ hedonic perceptions of users. We have first tried to predict the average eudaimonic/hedonic qualities for movies. However, for future work, we can try to model users and predict for each

Human-Computer Interaction Slovenia 2021, November 11, 2021, Koper, Slovenia
elham.motamedi@famnit.upr.si (E. Motamedi); marko.tkalcic@famnit.upr.si (M. Tkalcic)
0000-0003-0127-4997 (E. Motamedi); 0000-0002-0831-5512 (M. Tkalcic)

user what would be the value of the eudaimonic/ hedonic perception of a multimedia item. As a first step towards designing eudaimonia/hedonia-based recommender systems, we need to be able to label movies with their hedonic and eudaimonic qualities. This paper presents the preliminary results of a prediction model that marks movies using their subtitles.

2. Related Work

Multimedia items can be described with some characteristics. Typical movie characteristics that have been used extensively in recent studies are content-centric such as title, genre, director, actor and length. These characteristics are only related to the item and do not consider how the users perceive the items. The characteristics of the multimedia items can be calculated from text, audio, visual modalities or any combinations of these modalities. Deldjoo et al. [6] have proposed a content-based movie recommender system using different sets of such content-centric features. The weights of the features have been learned based on the behaviour of the users. They have deployed their recommender system for 13 months and have compared the results in terms of hit count using different feature sets. The features used in this study have included actors, directors, genre, and keywords that are all related to the content of the item and not to users' perceptions. In another study, Deldjoo et al. [7] have computed aesthetic low-level features from visual information of movie trailers and have exploited these features in recommender systems. They have shown that their proposed system has outperformed some base recommender systems using common high-level features such as genre. Chang and Ki [8] have proposed a new theoretical framework to predict the theatrical movie's success. They have found that different features, including length of run, objective elements, sequel actor, budget, genre and release periods, were significantly related to the box office performance.

Many studies have been using content-centric items' characteristics in different applications, including recommender systems. However, in recent years, there has been a growing interest in the new category of characteristics. Researchers have started to study characteristics of the items that depend on the users' perceptions of the items. One of the widely used examples is induced emotions. Induced emotion is the characteristic of the item and is based on which emotion it induces in users. Yang et al. [9] have manipulated the surgical images and have exposed users to the original and manipulated images. They have measured the users' emotional reactions and have found that the negative emotional reactions caused by the original images were higher than the manipulated images. Manipulating the emotional design of different multimedia learning materials and exploring the effect on other parameters such as learning outcomes and the users' cognitive process have also been studied in the literature [10, 11].

Other user-centric characteristics, such as induced stress levels and users' eudaimonic/ hedonic perceptions, have gotten less attention. Eudaimonic/hedonic qualities are concepts that are mostly studied in the psychology domain. Some works have introduced these concepts in the domain of recommender systems. Tkalčič and Ferwerda [4] have adapted these concepts to the domain of movies and have given a score of the eudaimonic/hedonic qualities to each movie of the dataset. They have found two clusters of users: (i) pleasure seekers, who prefer movies with higher value of hedonic quality and (ii) meaning seekers, who prefer movies with a higher score of eudaimonic quality. Chu et al. [12] have proposed an audio recommender system that chooses the music based on the events found in the text messages of a user's phone. Their idea has relied on this assumption that the users' preference for meaning in music is based on what is happening to them individually. Further work is needed to explore the effectiveness of exploiting eudaimonic/ hedonic qualities in other applications including recommender systems.

3. Preliminary Analysis of the Dataset

We have used the dataset collected by Puc [13], containing data on 177 users. They have filled-in a questionnaire and provided more information about five movies they have watched from a pool of 30 movies. Users' information contains the following information: genre, education, age, genre preferences, personality traits, sophistication index factors, and eudaimonic-hedonic orientation. Users' perceptions from movies include eudaimonic-hedonic perceptions, preference score, and REIVO factors. REIVO stands for relative assessment of orientation to extrinsic motivation rather than intrinsic motivation. If one does an activity for the extrinsic advantages of the activity rather than inner satisfaction, that leads to a higher score of REIVO. The REIVO's main factors are: financial success, popularity, image, community feeling, affiliation, and self-acceptance. The first three factors appear with positive weight in the formula of REIVO as they are more related to extrinsic motivation, and the last three factors with negative weight. The questions related to assessing each factor have been adapted for the movie domains. Genre, education, age, and genre preferences have been asked explicitly. However, personality traits, sophistication index factors, and eudaimonic-hedonic perceptions have been calculated from users' responses to the respective questions.

The Big five personality model has been used for describing users' personality traits. The film sophistication index is composed of three main factors: 1) active engagement, 2) perceptual abilities, and 3) emotions. Active engagement assesses the degree of being actively engaged in movies (as an example, one question here has been *how often the person reads or searches on the internet for the things or activities related to movies*). The perceptual ability assesses the ability to judge and notice different concepts about movies by watching a movie. For example, if a person can notice the genre of the movie by watching a movie, that increases the score of perceptual ability. Emotion is more related to the degree of being emotionally induced by movies. Another piece of information is the eudaimonic-hedonic orientation of users. It captures how much a person seeks meaning in movies or to which degree a person likes to watch movies that are just entertaining. A movie gets a higher score of eudaimonic quality if it contains a complex or deep meaning, and it gets a higher score of hedonic quality if it is more entertaining. Eudaimonic-hedonic perception of a movie is how much the user judges a movie as eudaimonic or hedonic.

For each user, we have had a score of eudaimonic and hedonic quality. We have been interested in checking if there is any significant correlation between other users' characteristics and the eudaimonic/ hedonic quality. Among different genres that users have been interested in, some have correlated with eudaimonic/hedonic quality significantly. For drama preference, there has been a significant correlation of $+0.7$ and -0.7 with eudaimonic and hedonic attributes, respectively. The correlations with eudaimonic quality for action and comedy preferences have been -0.5 and -0.4 and with hedonic quality +0.6 and +0.7, respectively. The values of correlations are not surprising as we expected that users who are more into comedy and action genres are likely to be from users who are more interested in being entertained than going deep into the thoughts raised by a movie. We also conjecture that the users who prefer drama are more meaning-seekers. Therefore, we have expected to have a higher score of eudaimonic quality for them. The same goes for two other genres of biographies and documentaries. We have found a correlation of $+0.4$ and $+0.5$ with eudaimonic quality and -0.5 for hedonic quality for two genres of biographies and documentaries, respectively.

Among many personality models, the big five personality traits have been extensively used by psychology's researchers to analyse people's behaviours. The five personality traits are openness, conscientiousness, extraversion, agreeableness and neuroticism. We wanted to see if there is a high correlation between some of these factors and the eudaimonic/hedonic qualities. Based on the results, users with a higher score of openness traits have been among users with a higher value of eudaimonic quality. The correlation between eudaimonic and openness has been calculated as $+0.6$, and the correlation between hedonic and openness has been -0.4. The correlation between extraversion and eudaimonic quality has been -0.3, and the correlation with hedonic quality has been $+0.3$. The correlation between other traits in the Big five-factor trait model and eudaimonic hedonic quality has not been noticeable.

The sophistication index includes three factors. Emotion is one of the factors that highly correlates with eudaimonic/hedonic quality. The easier the movies can induce emotions in a person, the higher the value of emotion in the sophistication index. The correlation between emotion and eudaimonic quality is $+0.8$, and the correlation between emotion and hedonic quality is -0.6, which is noticeably high.

Self-acceptance and community feeling are two factors of REIVO that correlate to eudaimonic/hedonic perception qualities. Correlation among different factors of REIVO and eudaimonic/ hedonic qualities are demonstrated as a heat map in Figure 1.

4. Methodology and Results

We conjecture that the subtitles of the movies include information about the eudaimonic/hedonic qualities of the movies. Therefore, we have aimed at developing a model that predicts the average eudaimonic/hedonic scores of movies given by users from the movie subtitles. In order to achieve this goal, we have considered the prediction problem as regression and classification.

In the regression problem, the aim is to predict the value of eudaimonic/hedonic scores. In the classification problem, we try to identify the eudaimonic/hedonic class of the movies. We have defined the eudaimonic and hedonic classes of the movies: 1) If the movies' eudaimonic score is higher than the median and the hedonic score is less than the median, the movie falls into the eudaimonic category. 2) If the hedonic score is higher than the median, and the eudaimonic score is less than the median, it falls into the hedonic category. 3) All other movies that do not belong to the eudaimonic or hedonic classes are labelled as other movies.

The subtitles are first preprocessed by applying some common preprocessing techniques used in NLP, including stemming, lemmatization, tokenization and removing stop words. Then the features have been extracted from the preprocessed subtitles. We have used TF-IDF and FASTTEXT pretraining models for feature extractions [14]. The list of algorithms that we used

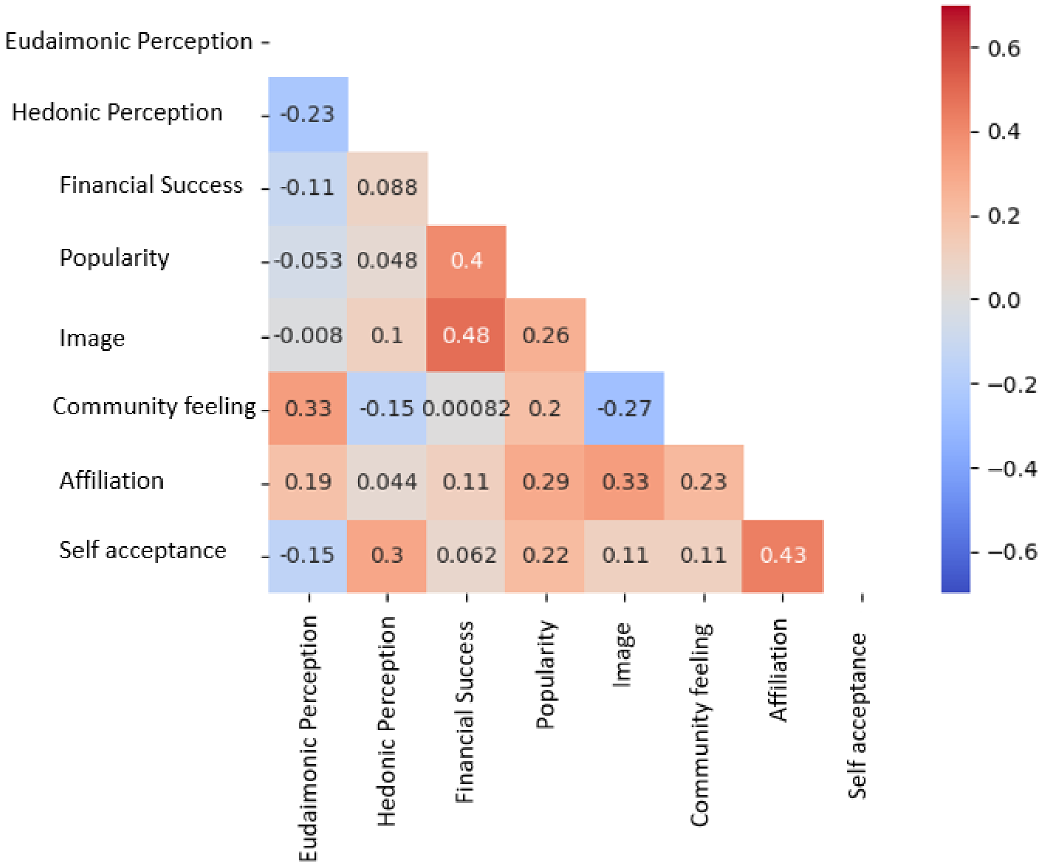

Figure 1: Correlation among REIVO factors of movies and eudaimonic/hedonic perceptions of users from the movies

and the pipeline of the work is presented in Figure 2.

We have used a nested k-fold cross-validation algorithm to evaluate the algorithms and compare the results with the base algorithms. In the regression algorithms, the predicted label by the base algorithm is the average of the labels in the training dataset. The base classification algorithm chooses the most frequent class as the new item's label.

The results of the regression algorithms are presented in Table 1. Decision tree, random forest regressor and XGboost have performed better in terms of root mean square error (RMSE) values in both eudaimonic and hedonic predictions. However, the mean absolute error (MAE) results are not always better than the base algorithms for the mentioned algorithms. We conjecture that the difference is related to the existence of some outliers in the dataset. The results of the classification algorithms are reported in Table 2. The accuracy and recall values of all

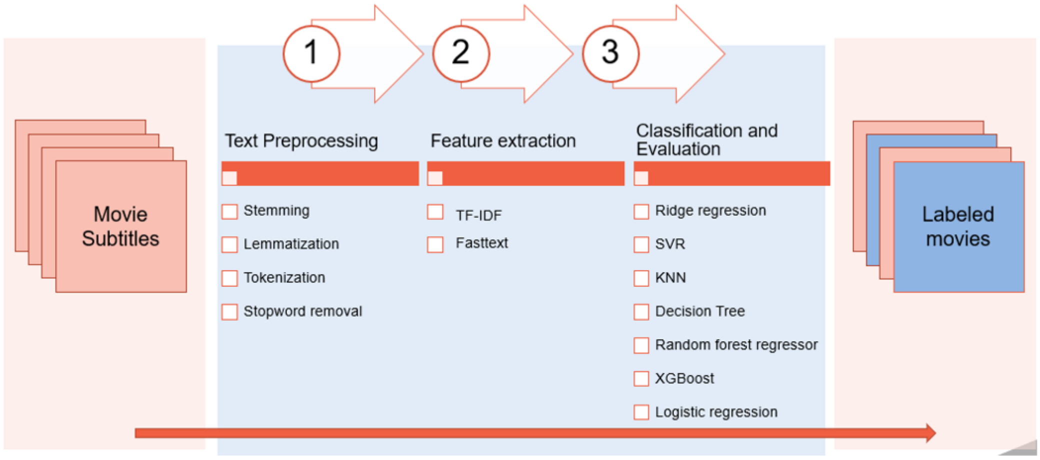

Figure 2: Methodology pipeline for predicting eudaimonic/hedonic label of movies from subtitles.

the classification algorithms are the same or better than the baseline. However, the decision tree, random forest and XGB classifiers are not always better based on precision, F1-score or ROC AUC values. The achieved results of this study may not be conclusive due to the limitation of having enough data. For future work, we have a plan to collect more data to improve the training models.

Table 1
Regression algorithms' results with k-nested cross-validation evaluation (10 outer and 3 inner folds)

Algorithms	Eudaimonic RMSE (sd)	Eudaimonic MAE (sd)	Hedonic RMSE (sd)	Hedonic MAE (sd)
Base Algorithm	1.26 (0.00)	1.06 (0.00)	1.34 (0.0)	1.20 (0.00)
Ridge Regression	1.29 (1.57)	1.08 (0.60)	1.29 (1.56)	1.08 (0.60)
SVR	1.31 (1.55)	1.07 (0.59)	1.29 (1.56)	1.07 (0.57)
KNN	1.29 (1.43)	1.05 (0.51)	1.29 (1.43)	1.05 (0.51)
Decision Tree	1.07 (0.81)	1.22 (0.61)	1.13 (0.56)	1.17 (0.59)
Random Forest	1.06 (0.73)	1.18 (0.58)	1.13 (0.64)	1.25 (0.34)
XGboost	1.19 (0.77)	1.22 (0.50)	1.16 (0.51)	1.20 (0.42)

5. Discussion and Future Work

In this work, we devised a model for predicting movies' eudaimonic/hedonic qualities from subtitles. To the best of our knowledge, there were no attempts to predict movies' eudaimonic/ hedonic qualities from subtitles prior to this study. The results of the regression and classification algorithms are, in most cases, better than the base algorithms. However, we assume that the results are not conclusive due to not having enough data. For future work, we are going to

Table 2
Classification algorithms' results with k-nested cross-validation evaluation (10 outer and 3 inner folds)

Algorithms	Accuracy (sd)	Precision (sd)	Recall (sd)	F1-Score (sd)	ROC AUC (sd)
Base Algorithm	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)	0.50 (0.00)
Ridge Classifier	0.62 (0.31)	0.60 (0.37)	0.85 (0.32)	0.63 (0.35)	0.78 (0.33)
SVM	0.58 (0.32)	0.58 (0.38)	0.85 (0.32)	0.62 (0.35)	0.57 (0.42)
KNN	0.65 (0.25)	0.58 (0.38)	0.85 (0.32)	0.63 (0.35)	0.65 (0.34)
Decision Tree	0.50 (0.25)	0.35 (0.39)	0.85 (0.32)	0.50 (0.35)	0.60 (0.23)
Random Forest	0.60 (0.32)	0.43 (0.32)	0.60 (0.44)	0.30 (0.38)	0.42 (0.39)
XGboost	0.50 (0.13)	0.35 (0.33)	0.55 (0.47)	0.40 (0.33)	0.72 (0.39)
Logistic Regression	0.62 (0.27)	0.55 (0.42)	0.75 (0.40)	0.57 (0.40)	0.57 (0.40)

collect more data to having more generalizable results. Moreover, we would like to model users to predict the eudaimonic/hedonic perception of each user for different movies. We are also interested in exploiting users' eudaimonic/hedonic perceptions in recommender systems for personalizing the recommendations.

References

[1] F. Ricci, L. Rokach, B. Shapira, Recommender Systems: Introduction and Challenges, in: Recommender Systems Handbook, volume 54, Springer US, 2015, pp. 1–34. URL: http://www.springerlink.com/index/10.1007/978-0-387-85820-3http://link.springer.com/10.1007/978-1-4899-7637-6_1. doi:10.1007/978-1-4899-7637-6_1.

[2] M. de Gemmis, P. Lops, C. Musto, F. Narducci, G. Semeraro, Semantics-Aware Content-Based Recommender Systems, in: Recommender Systems Handbook, Springer US, Boston, MA, 2015, pp. 119–159. URL: http://link.springer.com/10.1007/978-1-4899-7637-6_4. doi:10.1007/978-1-4899-7637-6_4.

[3] M. B. Oliver, A. A. Raney, Entertainment as Pleasurable and Meaningful: Identifying Hedonic and Eudaimonic Motivations for Entertainment Consumption, Journal of Communication 61 (2011) 984–1004. doi:10.1111/j.1460-2466.2011.01585.x.

[4] M. Tkalčič, B. Ferwerda, Eudaimonic Modeling of Moviegoers, in: Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization, ACM, New York, NY, USA, 2018, pp. 163–167. URL: http://dl.acm.org/citation.cfm?doid=3209219.3209249https://dl.acm.org/doi/10.1145/3209219.3209249. doi:10.1145/3209219.3209249.

[5] A. Delle Fave, F. Massimini, M. Bassi, Hedonism and eudaimonism in positive psychology, in: Psychological selection and optimal experience across cultures, Springer, 2011, pp. 3–18.

[6] Y. Deldjoo, M. Elahi, P. Cremonesi, F. Garzotto, P. Piazzolla, M. Quadrana, Content-based video recommendation system based on stylistic visual features, Journal on Data Semantics 5 (2016) 99–113.

[7] Y. Deldjoo, M. Elahi, M. Quadrana, P. Cremonesi, Using visual features based on mpeg-7 and deep learning for movie recommendation, International journal of multimedia information retrieval 7 (2018) 207–219.

[8] B.-H. Chang, E.-J. Ki, Devising a practical model for predicting theatrical movie success: Focusing on the experience good property, Journal of Media Economics 18 (2005) 247–269.

[9] H. Yang, J. Han, K. Min, Eeg-based estimation on the reduction of negative emotions for illustrated surgical images, Sensors 20 (2020) 7103.

[10] L. Stark, R. Brünken, B. Park, Emotional text design in multimedia learning: A mixed-methods study using eye tracking, Computers & Education 120 (2018) 185–196.

[11] J. Li, C. Luo, Q. Zhang, R. Shadiev, Can emotional design really evoke emotion in multimedia learning?, International Journal of Educational Technology in Higher Education 17 (2020) 1–18.

[12] S. L. Chu, S. Brown, H. Park, B. Spornhauer, Towards personalized movie selection for wellness: Investigating event-inspired movies, International Journal of Human–Computer Interaction 36 (2020) 1514–1526.

[13] E. Puc, Movie Recommender System – Psychological Constructs and General Movie Sophistication or Movie Likeability Score and Movie Genre, Ph.D. thesis, University of Primorska, 2021.

[14] E. Grave, P. Bojanowski, P. Gupta, A. Joulin, T. Mikolov, Learning word vectors for 157 languages, in: Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018), 2018.