Knowledge-Based Systems 270 (2023) 110539

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Automatic assignment of moral foundations to movies by word embedding

Carlos González-Santos [a], Miguel A. Vega-Rodríguez [b], Carlos J. Pérez [a,*], Joaquín M. López-Muñoz [c], Iñaki Martínez-Sarriegui [c,d]

[a] Departamento de Matemáticas, Universidad de Extremadura, Campus Universitario s/n, Cáceres, 10003, Spain
[b] Departamento de Tecnología de Computadores y Comunicaciones, Universidad de Extremadura, Campus Universitario s/n, Cáceres, 10003, Spain
[c] Research & Innovation Area, Consultora de Telecomunicaciones Optiva Media SL, Musgo 2 - Edificio Europa II, Madrid, 28023, Spain
[d] Research and Development Department, Metadatol SL, Santa Cristina 3 - Edificio Garaje 2.0, Cáceres, 10195, Spain

ARTICLE INFO

Article history:
Received 4 November 2022
Received in revised form 15 March 2023
Accepted 4 April 2023
Available online 6 April 2023

Keywords:
Moral foundations
Natural language processing
Semantic similarity
Word embedding

ABSTRACT

Morality is a topic that people are increasingly concerned about. Morality is observed and measured during public acts or when developing and consuming products, such as movies. The Moral Foundations Theory (MFT) was developed to rigorously perform these measurements with the support of the Moral Foundations Dictionary (MFD). In this paper, a Word Embedding-based Moral Foundation Assignment (WEMFA) approach has been designed, implemented, and applied to the movie domain for multiple assignment of moral foundations. WEMFA may use any dictionary, and it has been applied to a movie collection generated from movie synopses. A comparison between WEMFA and MoralStrength, the only approach found in the scientific literature, has been carried out. The proposed approach provided a percentage improvement of 41.7% with respect to the best version of MoralStrength, which uses an extension of the original MFD almost 10 times larger in number of terms. In addition, an extension of the original MFD (MFD24) has been built by adding 14 new moral foundations to the 10 original ones, enriching the moral context. WEMFA provided a mean accuracy of 78% with MFD24 despite the increment of the number of moral foundations. Besides, new extended dictionaries or even totally different ones can be used with WEMFA, since it does not need any training.

1. Introduction

Morality refers to certain conduct rules proposed by a society, or adopted by an individual, for their own behavior [1]. It is an essential feature of social life that has been at the core of many theories over the last century. It has given rise to investigations across multiple disciplines, including economics, anthropology, social psychology, comparative psychology, biology, and neuroscience [2].

The importance of morality has increased in the last several years. For example, a review of empirical research on psychology of morality led to 1,278 relevant scientific articles [3]. Five research themes were identified: moral behavior, moral reasoning, moral judgments, moral emotions, and moral self-views. This research confirms an increasing maturity of this area in terms of amount of research and relative impact.

Globalization has helped to standardize the definition of morality and classify behaviors, facts, and actions into right and wrong [4]. Indeed, this globalization process has led to a global concern for the conflicts present all over the world, like human rights, environmental, economics, and cultural issues, among others [5].

The moral foundations are the basis of any culture, which can vary among cultures according to the way they build their moral foundations. The main theory on morality is the Moral Foundations Theory (MFT) [6]. It is a theoretical framework that establishes a series of innate moral foundations that can be observed in any society. These moral foundations are paired as positive or negative attitudes: Care/Harm, Fairness/Cheating, Loyalty/Betrayal, Authority/Subversion, and Purity/Degradation.

Nowadays, a lot of information is daily generated: news, conferences, papers, tweets, etc. Through these texts, analysts can measure and evaluate the actions or behaviors of the author. Theoretical approaches like MFT may help to categorize individuals or entities in the moral context. In addition, categorization processes can be automatized and accelerated by using Natural Language Processing (NLP) tools [7].

* Corresponding author.
E-mail addresses: carlosgs@unex.es (C. González-Santos), mavega@unex.es (M.A. Vega-Rodríguez), carper@unex.es (C.J. Pérez), joaquin.lopez@optivamedia.com (J.M. López-Muñoz), inaki.martinez@metadatol.com (I. Martínez-Sarriegui).

https://doi.org/10.1016/j.knosys.2023.110539
0950-7051/

The combination of morality and NLP has led to interesting applications, which have been designed and implemented through different approaches. Essential tasks such as the detection or quantification of moral foundations have been developed by using techniques, such as topic modeling [8] or a combination of word embedding and knowledge graphs [9]. Moral changes over time have been traced by using word count and word embedding [7] or just with word embedding [10]. Also, there are some specific applications in the scientific literature. For example, the analysis of group-based morality in acts of hate [11]; the identification and tracing of politicians on tweets [12]; the analysis of moral expressions in palliative care consultations [13]; or the generation and selection of jokes based on moral judgement [14].

Multimedia contents are prone to be analyzed in terms of morality, since they are a very usual entertainment and may influence on people's behavior. Thus, morality is present in popular movies [15]. Indeed, online content platforms have many available contents and users must be warned about the possible presence of harmful content or just be informed about the moral themes that appear or are treated in the content. Movies can be tagged manually, but an expert team is required to watch and analyze each movie, which takes an excessive consumption of time and economic resources. Instead, the task of assigning moral foundations to movies can be automatized and carried out in real time by using NLP tools. Regarding the scientific literature, the presence of moral conflicts in movie scripts was detected by combining social network analysis and NLP [16]. They considered the script of a movie to build a graph of the relationships between characters through the movie scenes. Then, the relationship graph is used to extract moral conflicts among the characters.

All previous approaches work under the MFT by using the Moral Foundations Dictionary (MFD) [17]. Each moral foundation is represented by a list of terms, so that they provide a semantic definition of their corresponding moral foundation. This dictionary has been extended by adding new terms to each moral foundation [18]. Also, the relevance of the old and new added terms was measured and quantified with respect to the moral foundations [19]. The relevance of these terms was specifically measured only with respect to their corresponding moral foundation, providing a method named MoralStrength. This method performs a binary classification of moral foundations on tweets by considering the MFD and other extensions [20]. As it was defined, the method can only be used to assign one moral foundation to each piece of information.

To the best of authors' knowledge, all the approaches in the scientific literature performed an assignment of one moral foundation per piece of information. Nevertheless, texts may involve more than one moral foundation, which are often assigned manually. Assignment methods should be designed to perform a multiple selection and do it in an automatic way. To cover this scientific gap, we propose a Word Embedding-based Moral Foundation Assignment (WEMFA) approach that will be applied to the movie domain. Word2vec and BERT have been considered as methods for word embeddings. In addition, the number of 10 moral foundations from the dictionary has been maintained even in the extended dictionaries from previous researches. However, new moral foundations could cover a greater number of moral themes. Thus, an extension of MFD in terms of new moral foundations and their descriptions has been built.

The main contributions of this work are:

- For the first time, the problem of automatic assignment of multiple moral foundations has been addressed in the movie domain.
- Two word embedding-based approaches have been specifically designed, implemented, and applied to solve the problem of automatic assignment of multiple moral foundations to movies.
- The proposed Word2vec-based approach outperformed the BERT-based approach and the MoralStrength method, which uses an extended dictionary 10 times larger than the original MFD used with the proposed approach.
- An extension of MFD has been built to cover a wider range of moral themes.
- Results show a good performance and that the proposed extended dictionary enriches moral foundation assignment without a significant loss of accuracy.

The rest of the paper is organized as follows. In Section 2, the methodology is explained, including the extension of the dictionary and the proposed assignment method based on semantic similarity. The datasets, the evaluation metrics, the experimental setting, and the experimental results are presented in Section 3. Section 4 shows a discussion about the theoretical and practical implications of this research. Conclusions are detailed in Section 5. Finally, the extended dictionary and their descriptions are presented in Appendix.

2. Methodology

In this section, the word embedding techniques are explained. Next, the original MFD and the proposed extension are presented. Finally, the assignment method of moral foundations to movies is detailed.

2.1. Word embedding

Word embedding has become an important technique in the field of NLP [21]. It consists of a vector representation of a corpus vocabulary by embedding all its terms. The vector representation allows to calculate the semantic distance between each pair of terms. In addition, vector operations can be performed to calculate semantic similarities among a group of terms, such as sentences or documents. Word embedding has been applied in different NLP problems, such as text classification [22], machine translation [23], or document summarization [24].

Word2vec[1] is one of the most used models based on word embedding [25]. It was developed to work with huge datasets, providing a good performance in terms of accuracy and computation time. Word2vec implements two different models: continuous bag-of-words, which predicts a word based on the context; and skip-gram, which predicts nearby words given the current word. Google chose the skip-gram model, due to its efficiency, to be trained on a dataset from the Google News [26]. The dataset contains about 100 billion words and 3 million different terms (individual words and phrases) can be identified. Results provided by vector operations are really good in terms on semantics. For example, "Madrid" - "Spain" + "France" = "Paris".

Word2vec has been applied in many fields, such as sentiment analysis [27], anomaly detection [28], and malware classification [29], among others. In the movie domain, Word2vec has been used to recommend movies [30], perform a sentiment analysis on movie reviews [31], and to predict movie genres by using movie plots [32]. With respect to moral foundations, Word2vec-based approaches have been used in several applications, such as analyzing the moral language use of politicians [33], detecting moral biases in news [34], or investigating the relationships between morality and judicial decision-making [35].

Bidirectional Encoder Representations from Transformers (BERT)[2] is a pre-trained language model that provides word

[1] https://code.google.com/archive/p/word2vec/

[2] https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html

Table 1
Extended dictionary MFD24, composed by the original moral foundations of MFD and 14 new ones. Moral foundations are paired into positive and negative attitudes.

Moral foundation	Positive	Negative
Original	Care	Harm
	Fairness	Cheating
	Loyalty	Betrayal
	Authority	Subversion
	Purity	Degradation
New	Feminism	Maleness
	Sustainability	Climate change
	Racial equality	Racism
	Peace	War
	Liberty	Oppression
	Animal right	Animal abuse
	Sexual diversity	Sexual discrimination

embeddings attending to context [36]. The bidirectionality is one of its main features, since it considers both left and right context simultaneously. BERT is applied to solve many NLP tasks, such as question answering [37], language inference [38], and document classification [39]. It consists of two main steps: pre-training and fine-tuning. The pre-training process on unlabeled data is performed only once for different NLP tasks, saving time and computational effort. Then, the parameters of the pre-trained model are fine-tuned by using labeled data for each specific task.

BERT has been used for fake news detection [40], medical code assignment [41], and catering [42], among other applications. Regarding the movie domain, BERT has been used to apply sentiment analysis on movie reviews [43], and to build recommender systems [44]. With respect to moral foundations, BERT has been used for identifying ideological bias in news [45], and for moral concerns in social networks [46].

In this work, Word2vec and BERT will be integrated as alternatives in the multiple assignment model that is designed, implemented, and applied to solve the problem of multiple assignment of moral foundations. A comparison between both word embedding methods is presented.

2.2. Dictionaries

A dictionary consists of a list of concepts, each one defined by several terms. The MFD[3] is the most used dictionary in the morality context. It consists of a vocabulary of 324 terms that define the 10 moral foundations of the MFT. Besides, there is an extra category with general moral concepts that has been discarded because it does not represent any specific moral foundation, so the vocabulary remains in 295 terms.

The MFD has been extended several times by only adding new terms to the existing 10 moral foundations. Nevertheless, adding new moral foundations may better represent some of the most important current moral issues. Based on semantic similarity, MFD24, an extension of the original MFD with 14 new moral foundations, is proposed. They are also paired into positive and negative attitudes. Table 1 represents the original and new moral foundations.

The new moral foundations have been populated following the average number of terms per moral foundation of MFD, i.e., 30 terms have been selected to define each of them. Both unigrams and bigrams have been considered. The extension process has been carried out in the following way. Word2vec was used to find the most related terms to each of the new moral foundations. The 30 terms with the closest semantic distance with respect to a moral foundation were retrieved. Then, it was determined what terms were truly related to their corresponding moral foundation. Very similar terms or repeated concepts were removed. Finally, the list of terms of each moral foundation was completed by manual annotations in a second step until reaching 30 terms for each one.

The extended dictionary, including the moral foundations and their descriptions, is available in Appendix and can be obtained digitally upon request to the authors.

2.3. Assignment method

Let $\{M_m\}_{m=1}^{M}$ be a collection of M movies. Each movie is represented by a set of N_{M_m} tags $(t_{M_m 1}, \ldots, t_{M_m N_{M_m}})$. Also, the original MFD or the proposed extension MFD24, are represented as a collection of F (10 or 24) moral foundations $\{F_f\}_{f=1}^{F}$. Each moral foundation F_f is represented by a set of N_{F_f} terms $(w_{F_f 1}, \ldots, w_{F_f N_{F_f}})$. The goal is to assign the X most related moral foundations to each movie.

The assignment task is individually carried out for each movie. The set of tags from the synopsis of a movie is previously preprocessed. Firstly, the semantic similarities between the movie tags and the terms of each moral foundation are calculated. Specifically, for each movie tag, its semantic similarities are calculated with respect to every term of a moral foundation by using word embedding. The semantic similarities are expressed in the range [0,1]. If a semantic similarity is lower than a threshold, $thold_{sim}$, then it is not taken into account, since a very low semantic similarity between a term and a movie tag may introduce noise into the model. All the semantic similarities above the threshold are added to provide a total semantic similarity between a movie tag and a moral foundation. The process is repeated with all the movie tags, obtaining a vector of semantic similarities for a fixed moral foundation. Next, the moral foundation assignment is performed. To provide a semantic similarity between a moral foundation and a movie, the semantic similarities of every movie tag with respect to the moral foundation are added. This step is repeated for every moral foundation, resulting in a complete representation of a movie in the moral context. Finally, the moral foundations are sorted in descending order and the first X moral foundations are assigned to the movie. Fig. 1 shows a scheme of the full assignment method for all movies.

3. Results

In this section, the dataset for the experiments is presented. Next, the evaluation metrics and the experimental setting are described. Finally, the experimental results are detailed.

3.1. Dataset

To the best of authors' knowledge, there are no golden datasets on which to apply the task of multiple assignment of moral foundations. Then, a dataset from the movie domain including a collection of tags from 3,413 movies is considered. Movie information comes from their synopses, which provide descriptive information about the scenes, landscapes, characters or the movie plot. Synopses have been retrieved from well-known and publicly available sources: IMDb[4], Rotten Tomatoes[5], and FilmAffinity[6]. A total of 41,089 different tags have been obtained from the movie collection by using YAKE!, one of the most used models

3 https://moralfoundations.org/wp-content/uploads/files/downloads/moral%20foundations%20dictionary.dic

4 https://www.imdb.com/
5 https://www.rottentomatoes.com/
6 https://www.filmaffinity.com/

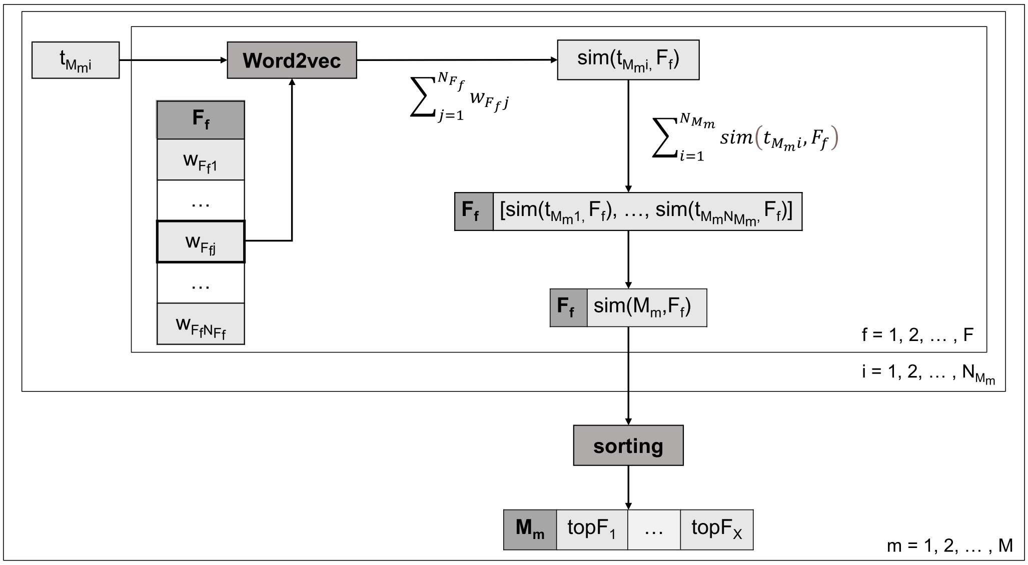

Fig. 1. Flowchart of the multiple assignment method of moral foundations to all movies.

for keyword extraction [47]. These tags have been automatically preprocessed in order to keep only those which could provide useful information about the movie and do not introduce noise into the model. The following two steps have been carried out:

1. Digit and symbol removal. Characters like digits, symbols or signs of punctuation are removed, so tags only keep their alphabetic form. If a tag results empty, it is removed from the tag list.
2. Stopword removal. Words such as articles or prepositions appear frequently and do not provide useful information, so they are removed from the tag. A tag is removed from the tag list if it is only composed by stopwords.

After the preprocessing, there are 34,891 different movie tags remaining. The movie dataset has been organized by establishing a relation between every movie and each of their tags, reaching a total of 109,664 movie-tag relations.

3.2. Evaluation metrics

The model performance has been evaluated by using the movie accuracy. The accuracy measures how many of the X top moral foundations are correctly assigned to a movie, i.e.:

$$Accuracy(M_m) = \frac{\text{number of correct topF}_x \text{ in } M_m}{X}. \quad (1)$$

The mean accuracy for all the movies is calculated as:

$$Accuracy_{mean} = \frac{\sum_{m=1}^{M} Accuracy(M_m)}{M}. \quad (2)$$

3.3. Experimental setting

A set of 20 popular movies from different genres and covering several moral issues have been selected to perform the experiments. All preprocessed data of these movies can be obtained digitally upon request to the authors.

The selection of a large number of moral foundations for the movie in relation to the total number of moral foundations in the dictionary might cause that low relevant moral foundations are assigned. Therefore, the number of top moral foundations to assign to a movie, X, is set at 20% of the total number of moral foundations. Thus, $X = 2$ for the original MFD (10 moral foundations), and $X = 5$ for MFD24 (24 moral foundations) are considered respectively.

The threshold for calculating semantic similarities, $thold_{sim}$, is set to 0.25 to ensure that term relationships are only considered when they are really similar. Other authors have used this same threshold value for solving problems related to semantic similarities [48,49].

The Word2vec-based approach (WEMFA-W2v) has been implemented and applied by using the Google pre-trained model. It has been trained by using 2 layers and 300 dimensions. In the case of the BERT-based approach (WEMFA-BERT), the base pre-trained model, which has been trained by using 12 layers and 768 dimensions, has been considered.

The MoralStrength method has been adapted to perform a multiple assignment of moral foundations, since it was originally designed for binary assignment [20]. This method has been applied by using three different versions of MFD, all of them with the 10 original moral foundations. The first one corresponds to the original MFD (295 terms) and allows a fair comparison with our approach. The second one is the extension proposed in [20] and contains 996 terms (Extended Dictionary 1, ED1). Finally, the third one is the version 1.1 of MoralStrength, composed of 2,845 terms (Extended Dictionary 2, ED2). Both extensions and the code are available online[7].

The experiments have been performed on a computer with an Intel Core i7-11700K CPU with 64 GB RAM, and Windows 10 as operating system. The assignment approaches have been implemented in Python 3.9 with the PyCharm 2021.1.3 IDE.

3.4. Experimental results

The first experiment consists of a model performance comparison. The MoralStrength method (the only approach found in the scientific literature), with the previously commented three dictionaries, and the two variants of the proposed WEMFA (WEMFA-W2v and WEMFA-BERT), with MFD, will be evaluated for assigning multiple moral foundations to a set of movies. Two moral foundations will be assigned out of the 10 original moral foundations for each movie. Model performance will be assessed by accuracy.

[7] https://github.com/oaraque/moral-foundations

Table 2
Number of moral foundations correctly assigned to movies out of 2 for WEMFA-W2v, WEMFA-BERT, and the MoralStrength method with the three versions of MFD.

Movie	# tags	WEMFA		MoralStrength		
		W2v-MFD	BERT-MFD	MS-MFD	MS-ED1	MS-ED2
2012	54	2	1	0	0	1
300: Rise of an Empire	52	2	2	1	1	1
A Single Man	38	1	1	0	0	0
Alexander	59	2	2	1	1	1
American Hustle	36	1	1	0	0	1
Avatar	67	2	2	0	0	1
Ben-Hur	52	2	1	0	0	1
Blood Diamond	30	2	1	2	2	2
Bohemian Rhapsody	48	1	0	0	0	1
Carmen & Lola	41	2	2	1	1	1
Cell 211	36	2	1	0	0	1
Fifty Shades of Gray	34	2	2	2	2	2
Finding Nemo	60	1	1	0	0	0
Intouchables	32	2	2	0	1	1
No Rest for the Wicked	54	2	1	2	2	2
Palm Trees in the Snow	43	2	2	2	2	2
Pride & Prejudice	45	2	2	2	2	2
The First Purgue	43	1	1	2	2	2
The Revenant	47	2	1	1	1	1
The Wolf of Wall Street	47	1	0	1	1	1
Average	45.90	1.70	1.30	0.85	0.90	1.20
$Accuracy_{mean}$	–	85.00%	65.00%	42.50%	45.00%	60.00%

Table 2 shows the results for the 20 selected movies. WEMFA-W2v provided a mean accuracy of 85.00%, with a mean of 1.70 (out of 2) moral foundations correctly assigned per movie. A mean accuracy of 65.00% was provided by WEMFA-BERT, with a mean of 1.30 (out of 2) moral foundations correctly assigned per movie. MoralStrength method (MS) with the three different dictionaries, MFD, ED1, and ED2, provided mean accuracies of 42.50%, 45.00%, and 60.00%, with a mean of 0.85, 0.90, and 1.20 (out of 2) moral foundations correctly assigned per movie, respectively. The more number of terms the dictionary has, the better it performs in accuracy terms. WEMFA-W2v and WEMFA-BERT made a perfect assignment in 14 and 8 (out of 20) movies, respectively. Meanwhile, the best MS approach, MS-ED2, made it perfectly only in 6 movies. Furthermore, whereas WEMFA-W2v always assigned correctly at least one moral foundation, WEMFA-BERT and MS-ED2 did not assign any correct moral foundation for 2 out of the 20 movies, and MS-MFD and MS-ED1 did not do it for 9 and 8 movies, respectively. In fact, MS-MFD and MS-ED1 sometimes assigned none or only one moral foundation instead of 2, thus being limited when detecting the presence of moral foundations.

WEMFA-W2v outperformed WEMFA-BERT by providing a 20% greater mean accuracy, which means an improvement percentage of 30.77%. Indeed, on a movie-by-movie comparison, WEMFA-W2v provided better or at least equal results than those from WEMFA-BERT. The assignment method was implemented to work at word-level, while BERT usually performs better while working at sentence-level, due to its context-based nature. From now on, WEMFA-W2v will be used in the next experiments and for notation simplicity we will refer to as WEMFA.

In order to illustrate the qualitative results obtained in a comparative context, the analysis of two movies (*300: Rise of an Empire* and *Carmen & Lola*) is presented. Table 3 shows the top moral foundations provided by WEMFA and MoralStrength approaches. WEMFA assigned correctly the 2 top moral foundations for both movies. MS-MFD only assigned correctly one moral foundation per movie, and it could not assign a second one for the movie *300: Rise of an Empire*. For this movie, MS-ED1 and MS-ED2 provided one correct moral foundation, whereas the other one was different (Betrayal and Fairness, respectively). For the movie *Carmen & Lola*, MS-ED1 and MS-ED2 only assigned correctly one moral foundation.

Table 3
Top moral foundations assigned by WEMFA and the MoralStrength approaches with several dictionaries for the movies *300: Rise of an Empire* and *Carmen & Lola*. Correct moral foundation assignments are in bold.

Movie	#Top	WEMFA	MS-MFD	MS-ED1	MS-ED2
300: Rise of an Empire	1	**Harm**	**Harm**	**Harm**	**Harm**
	2	**Degradation**	–	Betrayal	Fairness
Carmen & Lola	1	**Purity**	Loyalty	Loyalty	Loyalty
	2	Authority	**Authority**	**Authority**	**Purity**

The comparison between WEMFA and MoralStrength only was fair when the same dictionary was used, i.e., MFD. It is remarkable the efficiency of WEMFA that was superior to MoralStrength even when this last method was applied with dictionaries with more than 3 times (ED1) and almost 10 times (ED2) larger than MFD in number of terms. In addition, the weight of terms within the MoralStrength dictionaries are manually annotated. Thus, the use of new dictionaries requires an expensive and time-consuming process, whereas WEMFA works automatically with any dictionary (already existing or newly created).

The second experiment will evaluate the extended dictionary MFD24 and the behavior of WEMFA when increasing the number of moral foundations. In this case, 5 moral foundations (out of 24) were obtained as the top moral foundations for each movie.

Table 4 shows the results provided by WEMFA when using the original MFD and the extended dictionary MFD24. As it was presented before, WEMFA provided a mean accuracy of 85% when using the original MFD. In the case of MFD24, WEMFA provided a mean accuracy of 78%, with a mean of 3.90 (out of 5) correctly assigned top moral foundations per movie. 14 out of the 20 movies were assigned with an 80% or more of accuracy. Just 2 out of the 20 movies were assigned under a 50% of accuracy. The extended dictionary still leverages the original moral foundations, since 14 out of the 20 movies kept their correctly assigned top moral foundations from the original MFD.

The analysis of the movies *300: Rise of an Empire* and *Carmen & Lola* has also been considered in this case. Table 5 shows the top moral foundations assigned by WEMFA using both MFD and MFD24. WEMFA assigned correctly 2 and 5 top moral foundations when using MFD and MFD24, respectively. Furthermore, in the

Table 4
Results of the moral foundation assignment with the original MFD and the extended dictionary MFD24 for WEMFA.

Movie	# tags	WEMFA-MFD		WEMFA-MFD24	
		Top 2	Accuracy	Top 5	Accuracy
2012	54	2	100%	3	60%
300: Rise of an Empire	52	2	100%	5	100%
A Single Man	38	1	50%	4	80%
Alexander	59	2	100%	5	100%
American Hustle	36	1	50%	1	20%
Avatar	67	2	100%	4	80%
Ben-Hur	52	2	100%	5	100%
Blood Diamong	30	2	100%	4	80%
Bohemian Rhapsody	48	1	50%	3	60%
Carmen & Lola	41	2	100%	5	100%
Cell 211	36	2	100%	2	40%
Fifty Shades of Gray	34	2	100%	5	100%
Finding Nemo	60	1	50%	5	100%
Intouchables	32	2	100%	4	80%
No Rest for the Wicked	54	2	100%	3	60%
Palm Trees in the Snow	43	2	100%	4	80%
Pride & Prejudice	45	2	100%	5	100%
The First Purgue	43	1	50%	4	80%
The Revenant	47	2	100%	3	60%
The Wolf of Wall Street	47	1	50%	4	80%
Average	45.90	1.70	85%	3.90	78%

Table 5
Top moral foundations assigned by WEMFA, with MFD and MFD24, for the movies *300: Rise of an Empire* and *Carmen & Lola*. Correct moral foundation assignments are in bold.

Movie	#Top	WEMFA-MFD	WEMFA-MFD24
300: Rise of an Empire	1	**Harm**	**War**
	2	**Degradation**	**Oppression**
	3	–	**Harm**
	4	–	**Maleness**
	5	–	**Degradation**
Carmen & Lola	1	**Purity**	**Feminism**
	2	**Authority**	**Sexual diversity**
	3	–	**Purity**
	4	–	**Sexual discrimination**
	5	–	**Authority**

case of MFD24, it can be observed that the 2 correct moral foundations from the original MFD were retained while adding 3 more correct assignments from the new moral foundations.

The use of the extended dictionary did not cause an important loss of accuracy despite the increment of moral foundations from 10 to 24. Hence, MFD24 provides an enrichment of the moral foundations. Furthermore, the moral foundations have been defined by using general terms, so the dictionary can be applied to any domain. WEMFA demonstrated that performs well assigning multiple moral foundations, providing accurate results even with more than the double of moral foundations than those of MFD.

4. Discussion

Moral foundations are lately being more considered to analyze their presence in many kinds of content. However, the approaches available in the scientific literature assign those contents to only one moral foundation, whereas several of them may appear simultaneously in that content. An approach performing a multiple assignment of moral foundations would improve the results and enrich the solutions to this problem.

To the best of the authors' knowledge, WEMFA is the first approach assigning multiple moral foundations. WEMFA leverages word embedding to extract the semantic similarities of all terms with a certain meaning within a text and calculate how much each of the moral foundations is present on it. MoralStrength is the baseline approach for moral foundation assignment in a binary classification, so it has been adapted in this research work to allow multiple assignment. This statistical method is based on term matching, i.e., it requires that terms from a fixed dictionary must appear in the text to calculate the presence of some moral foundation. Thus, MoralStrength is strongly limited by a dictionary and may need a much larger dictionary to provide good enough results, whereas WEMFA performs well even with small dictionaries.

In addition, WEMFA leverages word embedding to completely automatize the multiple assignment of moral foundations. The automation also allows to use alternative dictionaries with different moral foundations and terms describing them with no necessary training. In the same way, WEMFA can process different datasets, leading the multiple assignment of moral foundations to other domains different from movies.

The extension of the MFD has determined, for the first time in the scientific literature, that new moral foundations can be added to the existing ones, providing richer solutions in terms of the number of moral foundations and without a significant loss of accuracy.

Since WEMFA is the first approach performing multiple assignment, there are no golden datasets to perform a comparison with the baseline approach. Despite this, it has been shown that the proposed method performs the best in a movie scenario. Both the movie dataset and the extended dictionary are digitally available upon request to the authors for future investigation purposes.

5. Conclusions

In this paper, the assignment of multiple moral foundations has been performed in the movie domain. A word embedding-based moral foundation assignment approach (WEMFA) has been designed, implemented, and applied to a movie dataset composed by tags extracted from the movie synopses. WEMFA has been implemented by using Word2vec and BERT.

Due to the lack of approaches in the scientific literature performing multiple assignment of moral foundations, MoralStrength [20] has been adapted and applied by using the original MFD and two more extended dictionaries. Both WEMFA approaches and MoralStrength have been evaluated and compared. The

Word2vec-based approach provided the greatest mean accuracy, becoming the best proposed approach. Despite the largest dictionary of MoralStrength was 10 times larger than the original MFD that used WEMFA, this approach performed the best providing a mean accuracy of 85%.

WEMFA may work automatically with any new dictionary because it does not need any manual annotation nor training. Then, it could be easily applied to other, even larger, dictionaries. In contrast, MoralStrength method uses manually annotated dictionaries, which makes it difficult and time-consuming to use new dictionaries.

The proposed MFD24 enriches the moral context by extending the number of moral foundations from 10 to 24. WEMFA kept accurate results when using MFD24, providing a mean accuracy of 78% despite the number of moral foundations was doubled. For future research, MFD24 could be applied to other specific domains to confirm the general purpose of this extended dictionary.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We are thankful to Optiva Media and Metadatol for providing the data necessary to conduct this study. This research has been supported by Ministry of Science and Innovation - Spain and State Research Agency - Spain (Projects PID2019-107299GB-I00 and PID2021-122209OB-C32 funded by MCIN/AEI/10.13039/501100011033), Junta de Extremadura - Spain (Projects IDA3-19-0001-3, GR21017, and GR21057), and European Union (European Regional Development Fund).

Appendix. Extended dictionary MFD24

See the extended dictionary MFD24 in Table A.6.

Table A.6
The extended dictionary MFD24 with its terms. Moral foundations are paired into positive and negative attitudes.

Foundation pair	Positive	Negative
Care/Harm (original)	safe, peace, compassion, empath, sympathy, care, caring, protect, shield, shelter, amity, secur, benefit, defen, guard, preserve	harm, suffer, war, wars, warlike, warring, fight, violence, hurt, kill, kills, killer, killed, killing, endanger, cruel, brutal, abuse, damag, ruin, ravage, detriment, crush, attack, annihilate, destroy, stomp, abandon, spurn, impair, exploit, exploits, exploited, exploiting, wound
Fairness/Cheating (original)	fair, fairly, fairness, fairminded, fairplay, equal, justice, justness, justify, reciprocal, impartial, egalitarian, rights, equity, evenness, equivalent, unbias, tolerant, equable, balance, homologous, unprejudiced, reasonable, constant, honest	unfair, unequal, bias, unjust, injust, bigot, discriminate, disproportion, inequitable, prejudice, dishonest, unscrupulous, dissociate, preference, favoritism, segregate, exclusion, exclude
Loyalty/Betrayal (original)	segregate, together, nation, homeland, family, families, familial, group, loyal, patriot, communal, commune, communit, communis, comrade, cadre, collective, joint, unison, unite, fellow, guild, solidarity, devotion, member, clique, cohort, ally, insider	abandon, foreign, enemy, betray, treason, traitor, treacherous, disloyal, individual, apostasy, apostate, deserted, deserter, deserting, deceive, jilt, imposter, miscreant, spy, sequester, renegade, terrorist, immigra
Authority/Subversion (original)	preserve, loyal, obey, obedient, duty, law, lawful, legal, dutifully, honor, respect, respectful, respected, respects, order, father, mother, motherly, mothering, mothers, tradition, hierarch, authority, permit, permission, status, rank, leader, class, bourgeoisie, caste, position, compliance, command, supremacy, control, submission, allegiance, serve, abide, deference, defer, revere, venerate, comply	betray, treason, traitor, treacherous, disloyal, apostasy, apostate, deserted, deserter, deserting, defiance, rebel, dissent, subvert, disrespect, disobey, sedition, agitate, insubordinate, illegal, lawless, insurgent, mutinous, defy, dissident, unfaithful, alienate, defector, heretic, nonconformist, oppose, protest, refuse, denounce, remonstrate, riot, obstruct
Purity/Degradation (original)	preserve, piety, pious, purity, pure, clean, sterile, sacred, chaste, holy, holiness, saint, wholesome, celibate, abstention, virgin, virgins, virginity, virginal, austerity, integrity, modesty, abstinent, abstemiousness, upright, limpid, unadulterated, maiden, virtuous, refined, decent, immaculate, innocent, pristine, church	ruin, exploit, exploits, exploited, exploiting, apostasy, apostate, heretic, disgust, depraved, disease, unclean, contagion, indecent, sin, sinful, sinner, sins, sinned, sinning, slut, whore, dirt, impiety, impious, profane, gross, repulse, sick, promiscuous, lewd, adultery, debauchery, defile, tramp, prostitute, unchaste, intemperate, wanton, profligate, filth, trashy, obscene, lax, taint, stain, tarnish, debase, desecrate, wicked, blemish, exploitation, pervert, wretched
Feminism/Maleness (new)	feminism, feminist, womanist, women libbers, postfeminist, femininity, womanhood, femaleness, womanliness, women liberation, female emancipation, sisterhood, women suffrage, empowerment, women, women rights, maternity, intersectionality, resistance, glass ceiling, women liberationist, libber, suffragette, reignite, self expression, girl power, female, revolution, movement, feminine	maleness, masculinity, manliness, effeminacy, animality, physical aggression, psychological aggression, masculine, selfhood, whiteness, misandry, anthropocentrism, masculinization, sluttiness, aestheticism, agelessness, manhood, virility, roughness, sex, patriarchy, sexism, misogyny, male gaze, verbal aggression, objectification, toughness, aggresive, boyhood, toxic masculinity

(continued on next page)

Table A.6 (*continued*).

Foundation pair	Positive	Negative
Sustainability/Climate change (new)	sustainability, sustainable, environmental, ecology, greening, eco, biodiversity, carbon neutrality, environmentally friendly, eco friendly, sustainably, sustainable forestry, renewable, ecological, efficient, natural, zero waste, low impact, organic, clean, preservation, conservation, biodegradable, regeneration, recycling, nonpolluting, reusable, environmental consciousness, environmental awareness, mitigation	climate change, global warming, climate, greenhouse gas, carbon emission, deforestation, desertification, resource depletion, climate crisis, volcanic eruption, adversity, carbon dioxide, emission, carbon footprint, drought, pollution, extinction, solar radiation, methane, climate emergency, wildfire, polar vortex, sea level rise, fossil fuel, flooding, hurricanes, ocean acidification, melting ice, heatwave, ozone layer
Racial equality/Racism (new)	equality, civil rights, racial harmony, emancipation, desegregation, nonviolence, desegregating, interracial, racial integration, egalite, ethnic equality, racial identity, racial tolerance, antiracism, equal opportunity, impartiality, inclusion, respect, integration, racial diversity, equal rights, equity, solidarity, social tolerance, equal treatment, unity, cultural understanding, racial solidarity, ethnic friendship, ethnic peace	homophobia, racism, bigotry, racist, racial discrimination, racial intolerance, racial prejudice, overt racism, sexism, xenophobia, classism, islamophobia, racial tension, stereotyping, bias, antisemitism, white supremacy, racial profiling, racial hostility, color bar, racialism, apartheid, segregation, sectarianism, separatism, prejudice, intolerance, narrow mind, dogmatism, partiality
Peace/War (new)	peace accord, peaceful coexistence, peacemaking, ceasefire, truce, reconciliation, disarmament, unity, amity, powersharing, tranquility, order, armistice, conciliation, harmony, security, treaty, calm, accord, concord, goodwill, friendship, cessation, neutrality, pacification, pacifism, unanimity, union, peacefulness, peaceable	war, invasion, occupation, conflict, hostility, warfare, military, trench, campaign, contend, combat, fight, struggle, bloodshed, contest, battle, enmity, enemy, murder, disagree, kill, shoot, army, crusade, soldier, weapon, terrorism, cold war, world war, ammunition
Liberty/Oppression (new)	freedom, liberty, democracy, unalienable rights, civil liberties, religious toleration, equality, democratic ideals, independence, responsibility, rightness, autonomy, civil rights, human rights, decision, emancipation, liberation, opportunity, determination, sovereignty, release, freeness, immunity, self rule, redemption, manumission, enfranchisement, deliverance, autarchy, leisure	repression, oppression, tyranny, subjugation, oppressor, injustice, colonialism, persecution, despotism, enslavement, marginalization, imperialism, dehumanization, patriarchy, totalitarianism, domination, depression, abuse, maltreatment, suppression, subjection, exploitation, brutality, coercion, cruelty, dictatorship, suffering, autocracy, harshness, calamity
Animal right/Animal abuse (new)	animal, animal welfare, animal protection organizations, exotic felines, domesticated animals, trap feral cats, disreputable breeders, animal care, animal protection, animal lover, cruelty free, protection, responsible, vaccination, habitat, compassion, liberation, health care, biosecurity, captivity, husbandry, veterinary care, wellbeing, adoption, microchip, natural habitat, animal control, pet care, legislation, medicine	animals, cruelty, abuser, mistreatment, animal overpopulation, neglect, maltreatment, distemper outbreak, deemed unadoptable, euthanized humanely, shelter euthanizes, zoos circuses, animal experiment, animal testing, exploitation, factory farming, suffering, abandonment, injury, vulnerable, threat, extinction, endangered species, vivisection, barbaric practice, aggression, gene pharming, euthanasia, exhibition, cosmetic testing
Sexual diversity/Sexual discrimination (new)	sexual diversity, cultural diversity, sexuality, homosexual, interracial friendships, gender, premarital sexual, sexual orientation, gender identity, transpeople, multicultural, sexual intercourse, heterosexual, diversity, antidiscrimination policies, equality, erotic arousal, transgender, pansexuality, sexual identity, asexual, bisexual, sexual preference, lesbian, gay, lgbt, polysexual, transsexual, sexual nature, gender diversity	sexual discrimination, gender discrimination, sexual harassment, nonconsensual sexual, sexual intercourse, gender bias, individuals discrimination, harassment, sexual abuse, sexual misconduct, sexual assault, intercourse, sexuality, antigay discrimination, sexual exploitation, homophobia, transphobia, prostitution, sexual objectification, prejudice, essentialism, naturalization, biphobia, heterosexism, coercion, domination, victim, unwanted, disorder, marginalized

References

[1] B. Gert, J. Gert, The definition of morality, in: The Stanford Encyclopedia of Philosophy, Metaphysics Research Lab, Stanford University, 2020, https://plato.stanford.edu/archives/fall2020/entries/morality-definition/. (Last Accessed 15 March 2023).

[2] M. Killen, J.G. Smetana, Origins and development of morality, in: HandBook of Child Psychology and Developmental Science: Socioemotional Processes, John Wiley & Sons, New Jersey, USA, 2015, pp. 701–749, http://dx.doi.org/10.1002/9781118963418.childpsy317.

[3] N. Ellemers, J. Van der Toorn, Y. Paunov, R. Van Leeuwen, The psychology of morality: a review and analysis of empirical studies published from 1940 through 2017, Pers. Soc. Psychol. Rev. 23 (4) (2019) 332–366, http://dx.doi.org/10.1177/1088868318811759.

[4] J. McKenzie, Globalization as a context for moral development, in: The Oxford HandBook of Moral Development: An Interdisciplinary Perspective, 2020, pp. 663–682, http://dx.doi.org/10.1093/oxfordhb/9780190676049.013.38.

[5] A. Benawa, To be a good global citizen, in: Proceedings of the 3rd International Conference on Social Sciences, Laws, Arts and Humanities - Volume 1: BINUS-JIC, 2018, pp. 350–354, http://dx.doi.org/10.5220/0010007803500354.

[6] J. Haidt, J. Graham, When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize, Soc. Justice Res. 20 (1) (2007) 98–116, http://dx.doi.org/10.1007/s11211-007-0034-z.

[7] J. Garten, R. Boghrati, J. Hoover, K.M. Johnson, M. Dehghani, Morality between the lines: Detecting moral sentiment in text, in: Proceedings of IJCAI 2016 Workshop on Computational Modeling of Attitudes, 2016.

[8] R. Kaur, K. Sasahara, Quantifying moral foundations from various topics on Twitter conversations, in: 2016 IEEE International Conference on Big Data (Big Data), 2016, pp. 2505–2512, http://dx.doi.org/10.1109/BigData.2016.7840889.

[9] Y. Lin, J. Hoover, G. Portillo-Wightman, C. Park, M. Dehghani, H. Ji, Acquiring background knowledge to improve moral value prediction, in: 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM, 2018, pp. 552–559, http://dx.doi.org/10.1109/ASONAM.2018.8508244.

[10] A. Ramezani, Z. Zhu, F. Rudzicz, Y. Xu, An unsupervised framework for tracing textual sources of moral change, in: Findings of the Association for Computational Linguistics: EMNLP 2021, 2021, pp. 1215–1228, http://dx.doi.org/10.18653/v1/2021.findings-emnlp.105.

[11] J. Hoover, M. Atari, A.M. Davani, B. Kennedy, G. Portillo-Wightman, L. Yeh, D. Kogon, M. Dehghani, Bound in hatred: The role of group-based morality in acts of hate, 2019, http://dx.doi.org/10.31234/osf.io/359me, PsyArXiv 359 me.

[12] S. Roy, D. Goldwasser, Analysis of nuanced stances and sentiment towards entities of US politicians through the lens of moral foundation theory, in: Proceedings of the Ninth International Workshop on Natural Language Processing for Social Media, 2021, pp. 1–13, http://dx.doi.org/10.18653/v1/2021.socialnlp-1.1.

[13] E. van den Broek-Altenburg, R. Gramling, K. Gothard, M. Kroesen, C. Chorus, Using natural language processing to explore heterogeneity in moral terminology in palliative care consultations, BMC Palliat. Care 20 (1) (2021) 23, http://dx.doi.org/10.1186/s12904-021-00716-3.

[14] H. Yamane, Y. Mori, T. Harada, Humor meets morality: Joke generation based on moral judgement, Inf. Process. Manage. 58 (3) (2021) 102520, http://dx.doi.org/10.1016/j.ipm.2021.102520.

[15] F. Yusuf, I. Fachrunnisa, Good-life concepts of moral value in robin hood movie, Lang. Cic.: J. Lang Lit. 16 (1) (2021) 193–200, http://dx.doi.org/10.15294/lc.v16i1.29556.

[16] F.R. Hopp, J.T. Fisher, R. Weber, A graph-learning approach for detecting moral conflict in movie scripts, Media Commun. 8 (3) (2020) 164–180, http://dx.doi.org/10.17645/mac.v8i3.3155.

[17] J. Graham, J. Haidt, B.A. Nosek, Liberals and conservatives rely on different sets of moral foundations, J. Personal. Soc. Psychol. 96 (5) (2009) 1029–1046, http://dx.doi.org/10.1037/a0015141.

[18] R. Rezapour, S.H. Shah, J. Diesner, Enhancing the measurement of social effects by capturing morality, in: Proceedings of the Tenth Workshop on Computational Approaches To Subjectivity, Sentiment and Social Media Analysis, 2019, pp. 35–45, http://dx.doi.org/10.18653/v1/W19-1305.

[19] F.R. Hopp, J.T. Fisher, D. Cornell, R. Huskey, R. Weber, The extended moral foundations dictionary (eMFD): Development and applications of a crowd-sourced approach to extracting moral intuitions from text, Behav. Res. Methods 53 (1) (2021) 232–246, http://dx.doi.org/10.3758/s13428-020-01433-0.

[20] O. Araque, L. Gatti, K. Kalimeri, MoralStrength: Exploiting a moral lexicon and embedding similarity for moral foundations prediction, Knowl.-Based Syst. 191 (2020) 105184, http://dx.doi.org/10.1016/j.knosys.2019.105184.

[21] Satvika, V. Thada, J. Singh, A Primer on Word Embedding, in: Data Intelligence and Cognitive Informatics, Springer, Singapore, 2021, pp. 525–541, http://dx.doi.org/10.1007/978-981-15-8530-2_42.

[22] S. Selva Birunda, R. Kanniga Devi, A review on word embedding techniques for text classification, in: Innovative Data Communication Technologies and Application, Springer, Singapore, 2021, pp. 267–281, http://dx.doi.org/10.1007/978-981-15-9651-3_23.

[23] H. Sun, R. Wang, K. Chen, M. Utiyama, E. Sumita, T. Zhao, Unsupervised bilingual word embedding agreement for unsupervised neural machine translation, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 1235–1245, http://dx.doi.org/10.18653/v1/P19-1119.

[24] R. Rani, D.K. Lobiyal, A weighted word embedding based approach for extractive text summarization, Expert Syst. Appl. 186 (2021) 115867, http://dx.doi.org/10.1016/j.eswa.2021.115867.

[25] T. Mikolov, K. Chen, G.S. Corrado, J. Dean, Efficient estimation of word representations in vector space, in: International Conference on Learning Representations 2013, 2013, http://dx.doi.org/10.48550/arXiv.1301.3781.

[26] T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, in: Advances in Neural Information Processing Systems, Vol. 26, NIPS, 2013, http://dx.doi.org/10.48550/arXiv.1310.4546.

[27] S. Styawati, A. Nurkholis, A.A. Aldino, S. Samsugi, E. Suryati, R.P. Cahyono, Sentiment analysis on online transportation reviews using Word2Vec text embedding model feature extraction and support vector machine (SVM) algorithm, in: 2021 International Seminar on Machine Learning, Optimization, and Data Science, ISMODE, 2022, pp. 163–167, http://dx.doi.org/10.1109/ISMODE53584.2022.9742906.

[28] J. Wang, C. Zhao, S. He, Y. Gu, O. Alfarraj, A. Abugabah, LogUAD: Log unsupervised anomaly detection based on Word2Vec, Comput. Syst. Sci. Eng. 41 (3) (2022) 1207–1222, http://dx.doi.org/10.32604/csse.2022.022365.

[29] A. Chandak, W. Lee, M. Stamp, A comparison of Word2Vec, HMM2vec, and PCA2Vec for malware classification, in: Malware Analysis using Artificial Intelligence and Deep Learning, Springer, Cham, 2021, pp. 287–320, http://dx.doi.org/10.1007/978-3-030-62582-5_11.

[30] Y.C. Yoon, J.W. Lee, Movie recommendation using metadata based Word2Vec algorithm, in: International Conference on Platform Technology and Service (PlatCon), 2018, pp. 1–6, http://dx.doi.org/10.1109/PlatCon.2018.8472729.

[31] A. Khan, D. Majumdar, B. Mondal, Machine learning approach to sentiment analysis from movie reviews using Word2Vec, in: Proceedings of Research and Applications in Artificial Intelligence, Springer, Singapore, 2021, pp. 131–140, http://dx.doi.org/10.1007/{978}-981-16-1543-6_12.

[32] Q. Hoang, Predicting movie genres based on plot summaries, 2018, http://dx.doi.org/10.48550/arXiv.1801.04813, 1801.04813.

[33] S.-Y.N. Wang, Y. Inbar, Moral-language use by US political elites, Psychol. Sci. 32 (1) (2021) 14–26, http://dx.doi.org/10.1177/0956797620960397.

[34] U. Shahid, B. Di Eugenio, A. Rojecki, E. Zheleva, Detecting and understanding moral biases in news, in: Proceedings of the First Joint Workshop on Narrative UnderstandIng, Storylines, and Events, 2020, pp. 120–125, http://dx.doi.org/10.18653/v1/2020.nuse-1.15.

[35] C. Brock, Moral Values and Judicial Decision-Making (Bachelor's thesis), Harvard College, USA, 2021.

[36] J. Devlin, M.W. Chang, K. Lee, K. Toutanova, BERT, pre-training of deep bidirectional transformers for language understanding, in: Proceedings of NAACL-HLT, 2019, pp. 4171–4186, http://dx.doi.org/10.18653/v1/N19-1423.

[37] P. Do, T.H.V. Phan, Developing a BERT based triple classification model using knowledge graph embedding for question answering system, Appl. Intell. 52 (1) (2022) 636–651, http://dx.doi.org/10.1007/s10489-021-02460-w.

[38] N. Jiang, M.C. de Marneffe, Evaluating BERT for natural language inference: A case study on the CommitmentBank, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP, 2019, pp. 6086–6091, http://dx.doi.org/10.18653/v1/D19-1630.

[39] J. Kong, J. Wang, X. Zhang, Hierarchical BERT with an adaptive fine-tuning strategy for document classification, Knowl.-Based Syst. 238 (2022) 107872, http://dx.doi.org/10.1016/j.knosys.2021.107872.

[40] S. Kula, M. Choraś, R. Kozik, Application of the BERT-based architecture in fake news detection, in: 13th International Conference on Computational Intelligence in Security for Information Systems, CISIS 2020, Springer, Cham, 2020, pp. 239–249, http://dx.doi.org/10.1007/978-3-030-57805-3_23.

[41] S. Ji, M. Hölttä, P. Marttinen, Does the magic of BERT apply to medical code assignment? A quantitative study, Comput. Biol. Med. 139 (2021) 104998, http://dx.doi.org/10.1016/j.compbiomed.2021.104998.

[42] M.H. Syed, S.T. Chung, Menuner: Domain-adapted BERT based NER approach for a domain with limited dataset and its application to food menu domain, Appl. Sci. 11 (13) (2021) 6007, http://dx.doi.org/10.3390/app11136007.

[43] B. Gupta, P. Prakasam, T. Velmurugan, Integrated BERT embeddings, BiLSTM-BiGRU and 1-D CNN model for binary sentiment classification analysis of movie reviews, Multimedia Tools Appl. 81 (23) (2022) 33067–33086, http://dx.doi.org/10.1007/s11042-022-13155-w.

[44] G. Penha, C. Hauff, What does BERT know about books, movies and music? Probing BERT for conversational recommendation, in: Proceedings of the 14th ACM Conference on Recommender Systems, 2020, pp. 388–397, http://dx.doi.org/10.1145/3383313.3412249.

[45] N. Mokhberian, A. Abeliuk, P. Cummings, K. Lerman, Moral Framing and Ideological Bias of News, in 12th International Conference on Social Informatics, SocInfo 2020, Springer, Cham, 2020, pp. 206–219, http://dx.doi.org/10.1007/978-3-030-60975-7_16.

[46] B. Kennedy, M. Atari, A.M. Davani, J. Hoover, A. Omrani, J. Graham, M. Dehghani, Moral concerns are differentially observable in language, Cognition 212 (2021) 104696, http://dx.doi.org/10.1016/j.cognition.2021.104696.

[47] R. Campos, V. Mangaravite, A. Pasquali, A. Jorge, C. Nunes, A. Jatowt, YAKE! keyword extraction from single documents using multiple local features, Inform. Sci. 509 (2020) 257–289, http://dx.doi.org/10.1016/j.ins.2019.09.013.

[48] A. Ustun, B. Can, Unsupervised morphological segmentation using neural word embeddings, in: Statistical Language and Speech Processing: 4th International Conference, Springer, Cham, 2016, pp. 43–53, http://dx.doi.org/10.1007/978-3-319-45925-7_4.

[49] G. Deepak, N. Kumar, D. Santhanavijayan, A semantic approach for entity linking by diverse knowledge integration incorporating role-based chunking, Procedia Comput. Sci. 167 (2020) 737–746, http://dx.doi.org/10.1016/j.procs.2020.03.339.