# Movie Review Analysis: Emotion Analysis of IMDb Movie Reviews

**Conference Paper** · August 2016

DOI: 10.1109/ASONAM.2016.7752387



Kamil Topal
Balikesir University

2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

# Movie Review Analysis: Emotion Analysis of IMDb Movie Reviews

Kamil Topal
Electrical Engineering and Computer Science
Case Western Reserve University
Cleveland, Ohio 44106
Email: kxt147@case.edu

Gultekin Ozsoyoglu
Electrical Engineering and Computer Science
Case Western Reserve University
Cleveland, Ohio 44106
Email: tekin@case.edu

***Abstract*—Movie ratings and reviews at sites such as IMDb or Amazon are commonly used by moviegoers to decide which movie to watch or buy next. Currently, moviegoers base their decisions as to which movie to watch by looking at the ratings of movies as well as reading some of the reviews at IMDb or Amazon. This paper argues that there is a better way: reviewers movie scores and reviews can be analyzed with respect to their emotion content, aggregated and projected onto a movie, resulting in an emotion map for a movie. One can then make a decision on which movie to watch next by selecting those movies having emotion maps with certain emotion map patterns desirable for him/her. This paper is a first step towards the above-listed scenario.**

***Index Terms*—Emotion Analysis, Sentiment Analysis, Dimension Reduction, Emotion Visualization**

## I. Introduction

It is common knowledge that, usually, moviegoers, called users or reviewers in the rest of this paper, utilize movie ratings and reviews in selecting their next movie to see/watch. This is indeed the case for the authors of this paper. And, unfortunately, sometimes movie reviews and ratings do not help users make the right choices, as evidenced by their emotional feelings after watching the movie. This is perhaps because users desire a certain emotional state after watching a movie, which does not match the emotions evoked by the selected movie. And, this is indeed what happened to the second author of this paper, after going over IMDb [1] ratings and (some) reviews, and picking, rather quickly and unfortunately, to watch the extremely highly-rated and indeed excellent Oscar-winning movie *The Revenant*–simply because the emotions evoked by the movie on him did not match his desired emotional state (as characterized by Cambria et als four emotional dimensions *pleasantness, attention, sensitivity and aptitude* [2]).

Clearly, users reviews and ratings for a movie are strongly tied to their emotions evoked by the movie. This paper argues that

1) It may be useful for users in their decision-making process to choose the next movie to watch if a movie also comes with an (expected) emotion signature or an emotion map, and,
2) Towards goal 1, we can build automated software tools that (i) analyze movie reviews and ratings, and (ii) provide an emotional signature for a movie as evidenced from the reviews and ratings of that movie.
3) Clearly, once emotion maps for all movies are at hand, if a user perhaps submits his desired emotion state and, possibly, the desired genre of the movie, it is easy to build a personalized movie recommender system for each user.

This paper is a first step for goals 1 and 2 (but not for goal 3, due to space limitations), and makes an attempt to analyze the relationships between

- Users ratings for a movie and their emotions evoked by watching the movie as evidenced in their movie reviews and ratings, and
- Movies genres and users emotional responses from their movie reviews.

For our experimental evaluation, we used movie reviews from IMDb, the worlds most popular content source for movie, TV and celebrity content [1] with reviews for more than 3.5 million movies. IMDb members provide reviews and usefulness scores for other reviews. However, the overall rating of a movie is calculated by IMDbs own rating algorithm, which, in turn, is based on reviewers scores [3].

Our movie dataset contains 157,344 reviews for 134 movies, where (i) most movies were very highly rated, (ii) a small number of movies were very lowly rated, and (i) selected moves spanned 19 genres/genre combinations, with each genre having at least 10 movies. Then, using the top 100 most useful-ranked reviews of each movie in our data set, we extracted for each movie an emotion map that captures reviewers emotions with respect to the four emotion dimensions of the Hourglass model of emotions of Cambria et al [2].

We have observed that (i) even for movies within the same genre, reviewers emotions may differ completely; that is, emotion maps of two very highly rated movies with the same genre labels may be drastically different. And, that (ii) emotion maps of very lowly rated (i.e., less than 3 out of 10) movies display minimal variation across the emotion scores of different reviewers; i.e., reviewers have very similar responses to a very low-rated movie.

Using the k-means clustering algorithm to cluster movies according to reviewers emotions per dimension, we have observed that

- All highly rated movies, which are those with ratings of 7 or above, have emotion maps with low or medium emotion levels, and rarely extreme emotion levels.
- Clustering allows effective genre-specific grouping of movies, meaning genre-specific movies have specific emotion map patterns as identified by the aggregated emotion maps of the groups that they belong.
- Genre coverage of some clusters are quite complete, meaning some clusters have significant numbers of movies from all genres.
- Removing emotion dimensions from clustering changes the clusters significantly.

Finally, we have redone the movie clustering with respect to all the reviews for a movie, not just the top 100 reviews of a movie. Our conclusion is that, within a cluster, the closest movies stay mostly the same regardless of the sizes of the emotion vectors used for cosine similarities. And, most similar movies still stay in the same cluster.

The rest of the paper is organized as follows. Section II summarizes the related work on emotion modeling, namely, the circumplex model of emotions and the hourglass of emotions model, and their datasets. Section III summarizes how emotions in user reviews are modeled and defines the notion of emotion vector of a movie (per emotion dimension) as well as the notion of emotion map of a movie.

## II. Related Work on Emotion Modeling

In 1980, J.A. Russell characterized emotions in two dimensions, namely, activation and pleasure, and proposed the circumplex model [4]. With his model, Stone provided a database containing lexicon of emotions, with more than 100 categories and 11,000 words [5].

Recently, Cambria et al proposed the hourglass model [2] which is a more advanced model than the circumplex model, and has four independent, but concomitant, dimensions, namely, pleasantness, attention, sensitivity, and aptitude (Figure 1). Each emotion dimension captures a different type of emotion:

(a) Pleasantness captures the users amusement level with interaction modalities,
(b) Attention captures interaction contents,
(c) Sensitivity captures the comfort level of the user with interaction dynamics, and
(d) Aptitude captures the users confidence in interaction benefits (Table I).

As seen in Figure 1 and Table I, each of the four dimensions of the hourglass model have six levels of activation, which collectively characterize the emotional state of an individual. Moreover, more complicated emotions can be seen in Table II by combining different dimensions. In total, from Tables I and II, each review has 8 different emotions by looking four independent dimensions and their pairwise combinations.

To form a dataset of phrases that capture different emotions, Cambria et al. first created the AffectNet dataset [6] by blending ConceptNet [7] and WordNet-Affect [8] datasets.

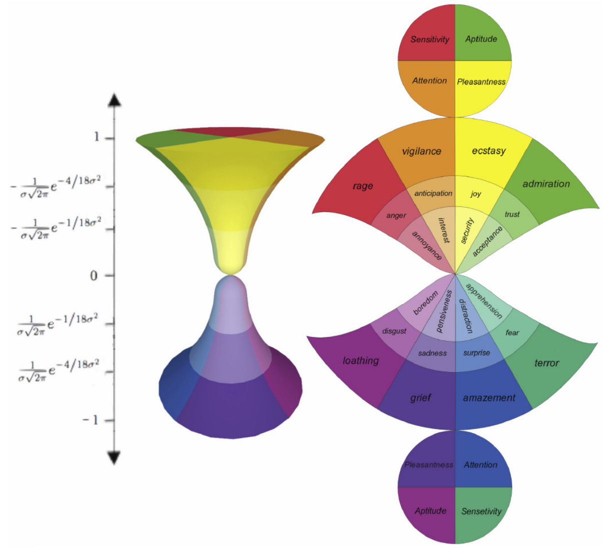


Fig. 1. HourGlass of Emotions Model [2]

TABLE I
HourGlass of Emotion Model

| | **Pleasantness** | **Attention** | **Sensitivity** | **Aptitude** |
|---|---|---|---|---|
| **+3** | ecstasy | vigilance | rage | admiration |
| **+2** | joy | anticipation | anger | trust |
| **+1** | serenity | interest | annoyance | acceptance |
| **0** | | | | |
| **-1** | pensiveness | distraction | apprehension | boredom |
| **-2** | sadness | surprise | fear | disgust |
| **-3** | grief | amazement | terror | loathing |

Then, they applied truncated singular value decomposition on AffectNet, and used dimension reduction on AffectNet by finding the best approximation. Finally they used the k-means clustering algorithm to cluster Sentic space to the HourGlass model. SenticNet 3.0 [9] database, which is publicly available, has more than 30,000 words and phrases that are already scored (in the range of $[-1, 1]$) for all dimensions. A snapshot of the SenticNet database is in Table III. In this paper, we use the hourglass model and the associated SenticNet database to classify each reviews emotion dimensions and the levels of user reviews.

## III. Modeling Emotion in Moviegoer Reviews

To represent the emotional landscape of each review of a movie, we use the SenticNet 3.0 [9] database. For each phrase in the database, there are four different scores (one for each

TABLE II
Compound emotions of second level

| | **Attntn >0** | **Attntn <0** | **Apttd >0** | **Apttd <0** |
|---|---|---|---|---|
| **Plsntnss >0** | optimism | frivolity | love | gloat |
| **Plsntnss <0** | frustration | disapproval | envy | remorse |
| **Snstvty >0** | aggressiveness | rejection | rivalry | contempt |
| **Snstvty <0** | anxiety | awe | submission | coercion |

TABLE III
A SNAPSHOT OF SENTICNET DATASET

| | **Pleasantness** | **Attention** | **Sensitivity** | **Aptitude** |
|---|---|---|---|---|
| accept | 0.972 | 0 | 0 | 0.894 |
| acceptability | -0.997 | 0 | 0.878 | -0.957 |
| acceptable | 0.258 | 0.357 | -0.088 | 0.351 |
| acceptance | 0 | 0 | 0 | 0.3 |
| acceptation | 0.577 | -0.77 | 0 | 0.517 |

of the four emotion dimensions) in the range [-1,1]. We map each review to the SenticNet database by identifying phrases in the review to a phrase in SenticNet. Since each comment may map to multiple phrases in the database, we then use dimension reduction for each review to get a single review score per emotion dimension. The scores of all reviews for a movie m per emotion dimension di forms the emotion vector of movie $m$ for emotion dimension $d_i$.

Let a movie $m$ have $n$ reviews, and the maximum number of SenticNet word instances in any review of movie $m$, duplicates included, be $k$. We create a $k$ by $n$ matrix to store all emotion scores of movie $m$ for emotion dimension, say, $d_i$. If a review has less than $k$ scores, we use bootstrapping [10] to $k$ dimensions for that review. Next, obtain an emotion dimension-specific review score for a review of the movie $m$, we reduce the matrix dimension from $k$ by $n$ to 1 by $n$ by using Singular Value Decomposition (SVD), as shown in Equation 1 and 2.

$$M_{d_i} = U_{d_i} \Sigma_{d_i} V^*_{d_i} \tag{1}$$

$M_{d_i}$ is a $k \times n$ original matrix for each dimension, $U_{d_i}$ is a $k \times k$ unitary matrix, $\Sigma_{d_i}$ is a $k \times n$ diagonal matrix with non-negative real number on the diagonal where the diagonal entries are singular values such as $\sigma_{11} > \sigma_{22} > \sigma_{33} > \cdots$, and $V^*_{d_i}$ is $n \times n$ unitary matrix.

$$E_{d_i} = (U^{(1)}_{d_i})^T M_{d_i} \tag{2}$$

Since $\sigma_{11}$ is the largest singular value, we take $U^{(1)}_{d_i}$ as first column of $U_{d_i}$ and transpose it from $k \times 1$ to $1 \times k$ and multiply by $M_{d_i}$ , then we get $E_{d_i}$ a $1 \times n$ vector, which contains a single review score for each review, which we call the emotion vector of movie $m$ for dimension $d_i$, where $d_i$ is one of dimensions pleasantness, attention, sensitivity, or aptitude.

An emotion map of a movie $m$ is simply the four emotion $E_{d_i}$ vectors of the movie, vertically stacked and represented as a heatmap that uses the same special red-green colorcoding scheme for all four emotion vectors.

## IV. EXPERIMENTAL EVALUATION AND RESULTS

### A. Dataset

Our dataset, collected using an IMDb API library [11], contains 157,344 reviews from (a) 116 highest-rated movies with large numbers of reviews at IMDb, and (b) 18 movies with low to very low reviewer scores. We also made sure that (i) our movie set had many genres, more specifically, 19 genres and genre combinations, and (ii) each genre category had at least 10 movies.

Note that (a) a movie may have multiple genre labels, (b) each review can be voted by other IMDb users for its usefulness, and (c) we collect individual user ratings of movies from reviews.

### B. Emotion Visualization

To visualize the aggregated per-dimension emotion score of a movies reviews, we use the movies emotion map which is really a special heatmap [12], a two dimensional graphical representation of data with colors instead of numbers. The reasons for using a heatmap are: (i) multiple dimensions can be seen in one figure, and (ii) different patterns can easily be distinguished from each other via clustering.

There are various color schemes for heatmaps [13]. Usually, with respect to color wheel [14], complementary colors are chosen. For our movie emotion maps, we choose the red-green scheme with six levels that represent an emotion on an emotion dimension. An example view of a movie emotion map is shown in Figure 2. The four emotion dimensions form the rows of the emotion map (i.e., placed horizontally), and each reviews scores per dimension form a column of the emotion map (i.e., placed vertically).

To save space, to construct a movie emotion map, we use the top 100 useful reviews of that movie within the reviews of the IMDb site. Moreover, to observe patterns within a movie emotion map, we hierarchically cluster reviews. For this reason, the original usefulness orders are altered in a heat map; e.g., the most useful comments place in a heatmap is not necessarily the first column.

### C. Results

As stated before, each review has 4 dimension scores. And, each reviews four emotion dimension scores are obtained as explained in Section III, by using the SenticNet database and Singular Value Decomposition (SVD).

*1) Emotion Heat Maps:* For each movie and a review, we compute the four emotion dimension scores, and create the corresponding movie emotion map. Figures 2, 3, 4, and 5 contain emotion maps of different movies.

**Observation 1**: The movie The Shawshank Redemption, from the most useful IMDb user reviews, evokes

- Mostly negative sensitivity dimension emotions, namely, apprehension and fear, and, only few users exhibit the terror emotion, and
- Mostly positive attention dimension emotions, i.e., interest and anticipation.
- The movie The Shawshank Redemption, rarely evokes the strongest emotion levels for all four dimensions (i.e., those close to +1 and -1).

**Observation 2**: Contrary to The Shawshank Redemption, the movie Godfather evokes

- Positive sensitivity dimension emotions, i.e., anger and annoyance;

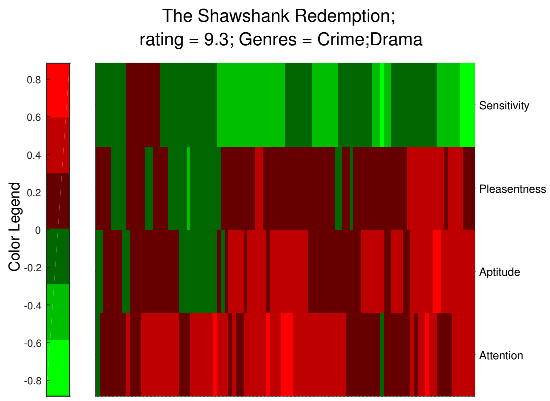


Fig. 2. Emotion Map of "The Shawshank Redemption".

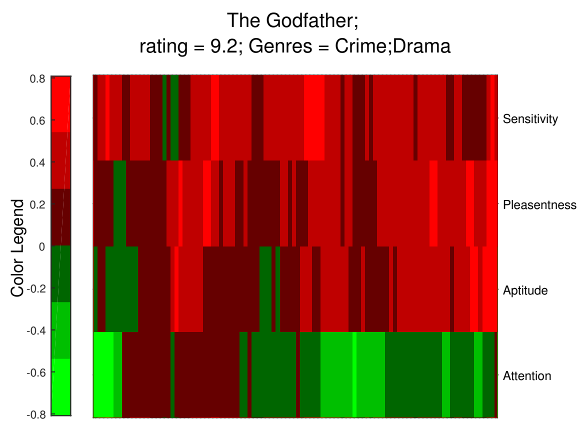


Fig. 3. Emotion Map of "The Godfather".

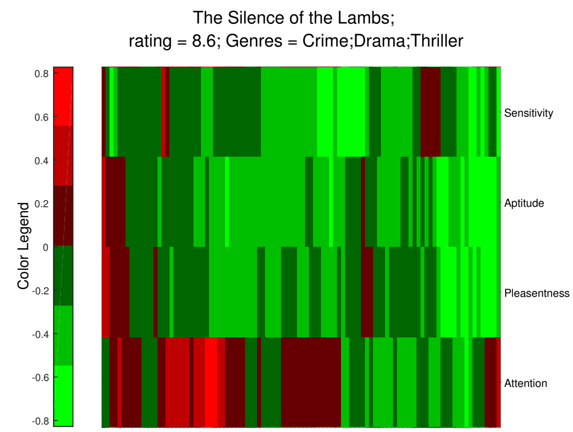


Fig. 4. Emotion Map of "The Silence of the Lambs".

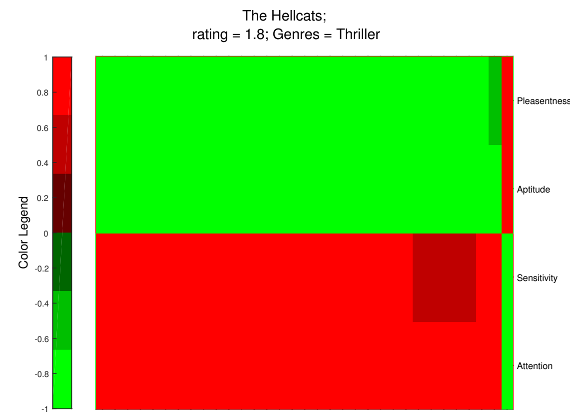


Fig. 5. Emotion Map of "The Hellcats".

- Negative attention dimension emotions, namely, distraction and surprise.
- Pleasantness and Aptitude dimensions look similar. Users exhibit joy and serenity; trust and acceptance emotions in different dimensions.
- Majority of reviews do not have polar emotions.

**Observation 3**: The movie The Silence of the Lambs evokes, for the majority of users, negative emotion levels in all dimensions.

**Observation 4**: The movie The Hellcats, which has a very low IMDb score, evokes very flat emotions and the same emotion score for all dimensions across all users.

**Observation 5**: Even for movies in the same genre, completely different reviewer emotions may be evoked. The movies The Shawshank Redemption and Godfather (Figures 2 and 3) belong to the same genre category, namely, Crime and Drama, and their IMDb movie ratings are very close, namely, 9.3 and 9.2 respectively. Yet, their emotion maps are completely different.

Note that emotions displayed in a review may reflect the failed expectations of the reviewers about the movie or the true emotions of the user, independent of his/her expectations. For example, a crime movie may have an anger score in a review because of its content and the movie has a good score; on the other hand, the same movie may have an anger score in another review simply because the expectations of the reviewer has not been met.

**Observation 6**: Most extremely low-scored movies evoke extreme emotions. The movie The Hellcats has a lower score and its emotion map has just one review pattern. For this movie, almost every reviewer shows extreme emotions on each dimension. From Pleasantness and Aptitude dimensions, there are remorse emotions among the reviews.

*2) Clustering Movies Based on Evoked Emotions :* There are six emotion levels within a dimension in the hourglass of emotion model. To interpret the polarity of emotions, we name -1 and +1 emotions as level 1 (the low level), -2 and +2 emotions as level 2 (the moderate level), and -3 and +3 emotions as level 3 (the extreme level) emotions based on Table I.

We use the k-means clustering algorithm to cluster movies according to reviewers emotions per dimension. First, we remove low-rated IMDb movies, which are those with ratings less than 7, from our dataset. This resulted in a movie data

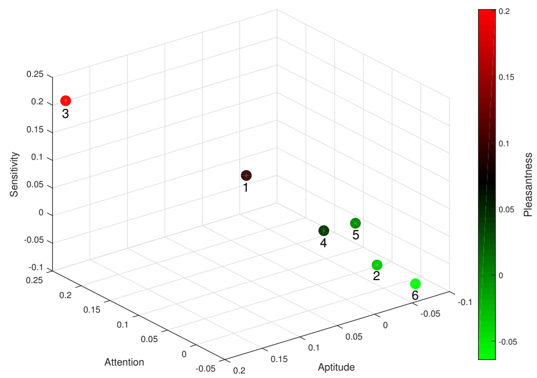


Fig. 6. k-means centroids. Since the data is four dimensional and the figure is three dimensional, the color within each circle indicates the score for the 4th dimension.

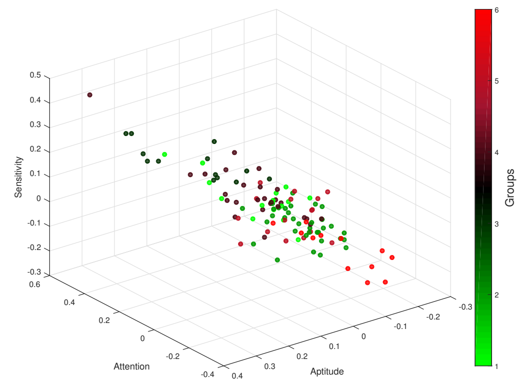


Fig. 7. Movies within each cluster. For visualization, the pleasantness dimension is removed from this figure. The colorbar shows the group (i.e., cluster) for each movie.

set of size 118. Again, for comparison purposes, the 100 most useful reviews per movie are used for clustering. There are 4 dimensions, and each movie has 400 reviewer scores across all four dimensions. After comparing the k-means algorithm with different k values, we have chosen k=6. Figure 6 shows the centroids of each of the six clusters, numbered 1 to 6 and the position of each centroid computed as the median of the cluster members scores visualized across three emotion dimensions, namely, sensitivity, attention, and aptitude. The median score value for the fourth dimension, namely, pleasantness, is represented by the color of each cluster centroid. Note that the extreme positions in each dimension are lost by taking the averages of dimension scores.

Next we investigate

- How close a movie is to its own clusters centroid,
- The distance between clusters, and
- The distance between movies from each other for each run.

**Observation 7**: Overall, centroids of some clusters are very close; so movies in these clusters may be similar even if they are in different clusters. Moreover, since all low-scored movies (i.e., those with scores less than 7) are removed, most observed emotions are in levels 1 or 2.

Figure 6 shows centroids of each group, and their distances to each other. Groups 3 and 6 are the farthest groups among all groups. On the other hand, groups 2, 4, and 5 are very close to each other. Distance from each movie can be seen in Figure 7. Table IV shows the closest 9 movies namely to their group centroids in the dataset.

Next we look into how genre distributions change among movies within a cluster. The total number of movies in the same genre in each group is divided by total number of movies of that genre. Figure 8 shows the results. As an example, 67% of all horror movies are in group 2, and there is no horror movie in groups 3, 5, and 6.

**Observation 8**: Movies belonging to certain genres are clearly distributed to specific groups, meaning genre-specific

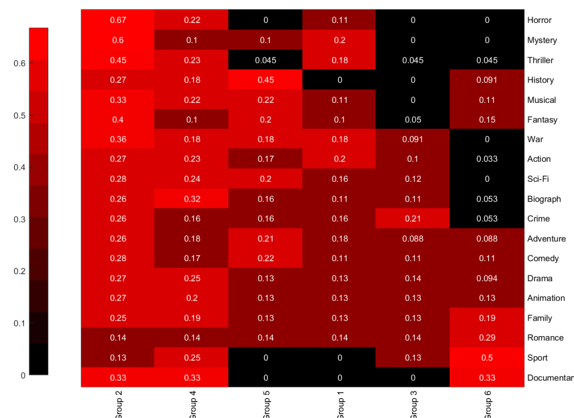


Fig. 8. Genre distribution of movies within the k-means (6-means) groups.

movies have specific emotion map patterns as identified by the aggregated emotion maps of the groups that they belong.

As examples, group 6 has 33% of documentaries and half of the sport movies in the dataset whereas group 3 does not contain any documentary movies and only 13% of sport movies. There is a specific Crime genre distribution as well.

**Observation 9**: Some groups have significant number of movies from all genres, e.g., group 2, which has movies belonging to almost every genre.

This may be because each movie may have multiple genre labels and these movies may evoke the same or similar emotions, regardless of the genre.

Next we investigate the relationships between clusters and specific emotions. So far, we have looked all dimensions together. However, users may be interested in, say, positive pleasantness movies. As also seen in Figures 2 and 3, movies may have common as well as distinct emotions, and, by concentrating on only specific dimensions or specific emotions, movies may belong to different or the same categories. To

TABLE IV
K-MEANS GROUPS, THEY ARE IN ASCENDING ORDER IN TERMS OF EUCLIDEAN DISTANCE TO BELONGING GROUP CENTROIDS

| **Group 1** | **Group 2** | **Group 3** |
|---|---|---|
| V for Vendetta | Saw | The Shawshank Redemption |
| The Godfather | The Avengers | 300 |
| World War Z | Pirates of the Caribbean: The Curse of the Black Pearl | The Dark Knight |
| Spider-Man | Psycho | Pulp Fiction |
| Inception | Avatar | Finding Nemo |
| Iron Man | Transformers | Interstellar |
| Inglourious Basterds | Se7en | Cars |
| American Beauty | Bowling for Columbine | Forrest Gump |
| The Usual Suspects | A Clockwork Orange | The Social Network |
| **Group 4** | **Group 5** | **Group 6** |
| The Lord of the Rings: The Fellowship of the Ring | The Green Mile | Million Dollar Baby |
| The Matrix | Braveheart | Titanic |
| Fight Club | Frozen | Rocky |
| The Dark Knight Rises | The Last Samurai | The Wizard of Oz |
| Iron Man 3 | 12 Years a Slave | How to Train Your Dragon |
| The Prestige | The Hunger Games | Kill Bill: Vol. 1 |
| WALLE | Schindler's List | The Silence of the Lambs |
| Gladiator | Gangs of New York | March of the Penguins |
| The Incredibles | District 9 | Zeitgeist |

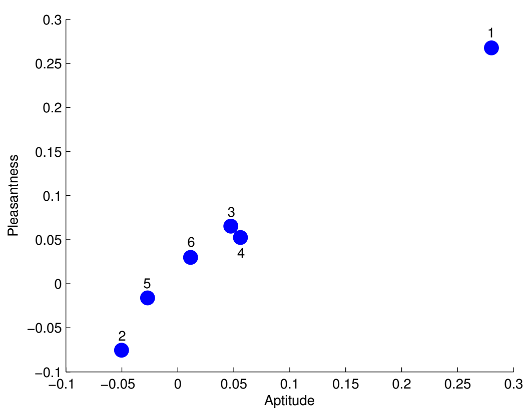


Fig. 9. Pleasantness vs Aptitude cluster centroids.

investigate this possibility, we ran the k-means algorithm with the same parameters except with only two dimensions. To save space, below we discuss the results of only one clustering run.

**Observation 10**: Running the k-means algorithm with only Aptitude and Pleasantness scores, we observe from Figure 9 that centroids of clusters are distinctively identified by some of the emotions love, gloat, envy, and remorse.

**Observation 11**: Reclustering movies after removing emotion dimensions selectively from clustering changes the clustering patterns significantly, indicating different automated movie recommendation possibilities based on changes in clustering dimensions.

TABLE V
SIMILAR AND DISSIMILAR MOVIES TO "V FOR VENDETTA"

| **Similar Movies** | **Dissimilar Movies** |
|---|---|
| Inception | The Shawshank Redemption |
| The Godfather | Pulp Fiction |
| Schindler's List | Interstellar |
| The Usual Suspects | The Silence of the Lambs |
| Iron Man | How to Train Your Dragon |

As examples, group 3 of Figure 6 and group 1 of Figure 9 are similar with respect to their Aptitude and Pleasantness scores. Group 3 in Figure 6 has 13 movies; on the other hand, group 1 of Figure 9 has 7 movies that exactly show up in group 3 of Figure 6. If specific emotions are desired, groups may change significantly.

*3) Emotion Vector Similarities of Movies:* Next, we change the dataset from the top 100 useful reviews to all reviews, and compare pairwise similarities between movies as represented by their emotion vectors (please see section III for the definition of an emotion vector of a movie). Note that the number of reviews for each movie may be different, and the size of the input for a movie as defined by the number of reviews of the movie needs to be the same for a pair of movies in order to calculate their pairwise similarity. So, we use the following approach to force them to be the same. When calculating the similarity between two movies with, say, $a$ and $b$ numbers of reviews, we first need to compute each movies emotion vector for the emotion dimension $d$. For the emotion vector computation, we always use $t = min(a, b)$ number of reviews from each of the two movies. As an example, the movie The Dark Knight Rises has 4,575 reviews, and the movie Inception has 2,777 reviews; thus, we take 2,777 most useful comments on The Dark Knight Rises and all of the reviews on Inception and calculate the cosine similarity of emotion vectors of the two movies for each of the four emotion dimensions. For emotion vectors, we take all of the dimensions, but each dimension can be taken independently or pairwise, as well.

One of the results is shown in Table V for the movie V for Vendetta which is the movie closest to group 1 centroid in Figure 6 and Table IV.

**Observation 12:** The closest movies stay mostly the same regardless of the sizes of the emotion vectors used for cosine similarities. And, most similar movies stay in the same group.

Observation 12 is useful for an initial starting point to a customized movie portfolio for a user. If a user writes a review and rate a movie with a good score; we calculate his emotions; then our emotion maps can recommend to him other movies based on his past ranking/emotion trend using other similar user reviews.

## V. CONCLUSIONS

We have introduced the notion of movie emotion maps for capturing the emotion content of movies via the emotions expressed in movie reviews. We have then analyzed the characteristic of a set of movie reviews in the IMDb database, towards the goal of providing movie recommender systems based on the emotion maps of movies and the desired emotion maps of moviegoers.


## Acknowledgment

This research is supported by a MEB scholarship for the first author from the goverment of Turkey and the US NIH grant R01HS020919-01. The conference travel funds for the presentation of this research by the first author are provided by the government of Turkey.


## References


[1] IMDb, "What is IMDb." [Online]. Available: http://www.imdb.com/pressroom/

[2] E. Cambria, A. Livingstone, and A. Hussain, "The hourglass of emotions," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7403 LNCS, 2012, pp. 144–157.

[3] IMDb, "The vote average for film." [Online]. Available: http://www.imdb.com/help/show_leaf?votes

[4] J. A. Russell, "A circumplex model of affect." *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980. [Online]. Available: http://psycnet.apa.org/journals/psp/39/6/1161

[5] P. Stone, "General Inquirer." [Online]. Available: http://www.wjh.harvard.edu/ inquirer/Home.html

[6] E. Cambria, A. Hussain, C. Havasi, and C. Eckl, "SenticSpace: Visualizing opinions and sentiments in a multi-dimensional vector space," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6279 LNAI, no. PART 4, pp. 385–393, 2010.

[7] C. Havasi, R. Speer, and J. B. Alonso, "ConceptNet 3: a Flexible, Multilingual Semantic Network for Common Sense Knowledge," *In Proceedings of Recent Advances in Natural Languges Processing 2007*, pp. 1–7, 2007. [Online]. Available: papers3://publication/uuid/154B5FA8-A3AB-4993-9833-B1C7340A9707

[8] C. Strapparava and A. Valitutti, "WordNet-Affect: an affective extension of WordNet," *Proceedings of the 4th International Conference on Language Resources and Evaluation*, pp. 1083–1086, 2004.

[9] E. Cambria, D. Olsher, and D. Rajagopal, "SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis," *Twenty-eighth AAAI conference on …*, pp. 1515–1521, 2014. [Online]. Available: http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8479

[10] D. a. Freedman, "Bootstrapping Regression Models," *The Annals of Statistics*, vol. 9, no. 6, pp. 1218–1228, 1981.

[11] "Python Library for IMDb json." [Online]. Available: https://github.com/richardasaurus/imdb-pie

[12] C. ??kuta, P. Bartu??n??k, and D. Svozil, "InCHlib - Interactive cluster heatmap for web applications," *Journal of Cheminformatics*, vol. 6, no. 1, 2014.

[13] C. Brewer, "Color Brewer." [Online]. Available: http://colorbrewer2.org/

[14] C. Parkhurst and R. L. Feller, "Who invented the color wheel?" *Color research and application*, vol. 7, no. 3, pp. 217–230, 1982.