The Application of Affective Measures in Text-Based Emotion Aware Recommender Systems

John Kalung Leung, Igor Griva, William G. Kennedy, Jason M. Kinser, Sohyun Park, and Seo Young Lee

Keywords: Emotion Aware Recommender Systems, Affective Computing, Users and Items Emotion Profiles, Text-Based Emotion Detection and Recognition, Affective Indices and Affective Index Indicators, Emotion Identification.

Abstract: This paper presents an innovative approach to address the problems researchers face in Emotion Aware Recommender Systems (EARS): the difficulty and cumbersome collecting voluminously good quality emotion-tagged datasets and an effective way to protect users’ emotional data privacy. Without enough good-quality emotion-tagged datasets, researchers cannot conduct repeatable affective computing research in EARS that generates personalized recommendations based on users’ emotional preferences. Similarly, if we fail to protect users’ emotional data privacy fully, users could resist engaging with EARS services. This paper introduced a method that detects affective features in subjective passages using the Generative Pre-trained Transformer Technology, forming the basis of the affective index and Affective Index Indicator (AII). Eliminate the need for users to build an affective feature detection mechanism. The paper advocates for a Separation of Responsibility approach where users protect their emotional profile data while EARS service providers refrain from retaining or storing it. Service providers can update users’ affective indices in memory without saving their privacy data, providing affective-aware recommendations without compromising user privacy. This paper offers a solution to the subjectivity and variability of emotions, data privacy concerns, and evaluation metrics and benchmarks, paving the way for future EARS research.

1 INTRODUCTION

The Emotion Aware Recommender System (EARS) aims to provide personalized recommendations by considering users’ affective preferences and opinions from similar users. However, researching EARS faces several challenges due to human emotions' subjective and variable nature (Qian et al., 2019). These challenges involve the development of accurate models for emotion detection, classification, and prediction, as well as collecting sufficient emotion-tagged datasets, which are hindered by the subjective and contextual aspects of emotions (Schedl...
et al., 2018). Additionally, the ethical considerations surrounding user privacy and data protection pose a significant challenge in EARS research (Bobadilla et al., 2013). In order to evaluate the effectiveness of EARS, we require robust evaluation metrics and benchmarks. However, the definition and measurement of the impact of emotions on user behavior and satisfaction are complex and multifaceted (Mohammad, 2016). Therefore, this paper proposes an innovative approach to address the following challenges faced by EARS researchers: the difficulty of collecting voluminous, high-quality emotion-tagged datasets to enhance affective computing and the need for an effective method to safeguard users' emotional data privacy.

In our envisioned world, affective metric signatures known as affective indices assign objects derived from subjective descriptions. Unless the subjective descriptions are revised, these affective indices remain static. However, users' affective indices are dynamic and evolve through their interactions and consumption of objects.

Our research introduces a novel approach to tackle EARS challenges and enhance their effectiveness. We address the difficulties associated with collecting sufficient high-quality emotion-tagged datasets and ensuring privacy, which impedes research in affective computing for personalized recommendations based on users' emotions.

We leverage Generative Pre-trained Transformer (GPT) technology to eliminate users' need to develop affective feature detection mechanisms. GPT effortlessly detects affective features in subjective passages.

We advocate employing the affective index and Affective Index Indicator (AII) as the foundation for detecting affective features and measuring emotions. We reckon that preserving the privacy of emotional data is crucial to prevent user resistance. We advocate for a Separation of Responsibility (SoR) approach, where users are responsible for protecting their emotional profile data while service providers refrain from storing it. The service providers can ensure efficient and personalized recommendations by updating users' affective indices in memory without compromising privacy.

Our research provides solutions to address subjectivity, variability of emotions, data privacy concerns, and evaluation metrics. These solutions contribute to the advancement of EARS research and its practical implementation.

## 2 RELATED WORKS

The main challenge in Text-based Emotion Aware Recommender Systems (EARS) research is obtaining high-quality, emotion-tagged datasets necessary for machine learning processing. To address this challenge, Guo (Guo, 2022) illustrated a deep learning-assisted semantic text analysis approach that involves defining emotion keywords, identifying data sources, developing a collection plan, cleaning and pre-processing data, and evaluating and refining the dataset. However, researchers still need more high-quality emotion-labeled datasets, which are required to train the emotion prediction model for classification. While standards Recent advances in transformer-based models, such as Generative Pre-trained Transformer (GPT) technology, have yet to establish benchmarking datasets for generating, PaLM, GPT-3, ChatGPT, BERT, ELMo, RoBERTa, and Transformer-XL offer promising new approaches to dataset collection (Ethayarajh, n.d.). These models have been trained on large amounts of text data to generate emotion labels (Kusal et al., 2022). They may provide valuable resources for researchers in this field.

### 2.1 Affective Tagged Datasets

EARS needs voluminously good quality emotional tags datasets for model training of making personalized recommendations. EARS requires Emotional tags to refer to labeling data, such as Ekman’s six basic emotions: happiness, sadness, anger, fear, surprise, and disgust (Humintell, 2020). These tags are crucial in developing EARS technology that can provision users’ and items’ emotional profiles. However, collecting accurate and sufficient emotional tags can be extremely challenging due to subjective and variable emotions (Russell, 2003), lack of standard tagging methods, time and resource intensity (Lo et al., 2017), privacy issues (Kompan et al., 2015), limited diversity (Yang et al., 2021), and contextual factors (Kanj et al., 2015). Therefore, collecting emotional tags for EARS research demands thoughtful planning and consideration of these factors (Mauss & Robinson, 2009a) in data features engineering.

### 2.2 Challenges of Affective Index Modeling in Personalized Recommender Systems

Recommender systems are crucial in modeling users' emotional profiles and item preferences. The
affective index, quantifying human emotion metrics using probabilistic values, has emerged as a valuable tool to achieve this goal (Acheampong et al., n.d.). In recent studies, Leung et al. (Leung et al., 2021) proposed the Affective Index Indicator (AII), which employs the Cosine Similarity metrics to generate a list of peer-wise numerical similarity values.

\[
\text{Inner}(x, y) = \sum_i x_i y_i = \langle x, y \rangle \\
\text{CosSim}(x, y) = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2} \sqrt{\sum_i y_i^2}} = \frac{\langle x, y \rangle}{||x|| ||y||} \tag{2}
\]

This computation enables the measurement of affective indices between an active user or item and its peers (Leung et al., 2020b), utilizing either the Nearest Neighbor algorithm (Keller et al., 1985) or alternative methodologies.

Over time, AII approaches have undergone significant advancements. Initially based on simple word lists (Kratzwald et al., 2018), they have evolved to incorporate more sophisticated techniques such as machine learning (Nasir et al., 2020) and lexicon-based methods (Hajek et al., 2020). Notably, recent research efforts have aimed to combine multiple techniques, resulting in a more robust AII (Naseem et al., 2020). However, the utilization of AII in recommender systems poses challenges regarding reliability, interpretability, and ethical considerations, particularly regarding privacy, stereotypes, and biases (Mauss & Robinson, 2009b).

Ongoing research focuses on developing more accurate and effective methods, identified by Leung et al. (Leung et al., 2020a), for measuring emotions in the text to address these challenges. These advancements can significantly enhance the relevancy and accuracy of personalized EARS. By leveraging the advances in measuring emotions, recommender systems can offer highly tailored recommendations, improving user satisfaction and engagement.

3 METHODOLOGY

3.1 Generate Emotion Tagged Data in EARS Research Through GPT

This paper proposes leveraging the extensive GPT database to obtain emotion-tagged data for EARS research through short prompting conversational dialogue. Researchers can query GPT for the affective indices of subjective texts, utilizing massive domain information when available. Using GPT in this way can easily extract affective indices from object descriptive passages and potentially become the standard method for gathering emotional labels.

Emotion Aware Recommender Systems (EARS) employ affective indices and Affective Index Indicators (AII) to measure emotions in users and items. AII calculates the similarity between a source object and target objects based on Ekman's six basic human emotions (Leung et al., 2020b). However, using AII poses challenges regarding reliability, interpretability, and avoiding biases. Further research is needed to enhance AII. Artificial intelligence language models like ChatGPT have shown promise in sentiment and affective analysis of a text. This study shows how the researchers prompt ChatGPT to build an affective index from the inspirational text.

To analyze sentiment in subjective passages, one must perform pre-processing, which involves noise removal, tokenization, and polarity assignment. Various sentiment analysis libraries and tools are available for this purpose. By aggregating the polarity scores, we can obtain an overall AII that we can utilize in downstream applications, such as user filtering and personalized recommendations.

ChatGPT can estimate the probabilistic values of basic emotions in subjective passages using sentiment analysis techniques and lexicons or machine learning models (Munn et al., 2023). The scores are normalized to create an affective index. While sentiment analysis is imperfect, our method provides a helpful guide for assessing emotional content (Lauriola et al., 2022). The AII and affective index offer valuable tools for recommendation systems and emotional filtering.

3.2 The Separation of Responsibility Framework Principles

We strongly advocate for protecting users' emotional data privacy through a Separation of Responsibility (SoR) framework. Here is how it works:

- Each object in the system is assigned an emotion ID (eID) based on its affective indices.
- The eID of an item object remains static, while a user's eID is dynamic, evolving with their interactions.
- Users are responsible for managing their eID under the SoR framework.
- Service Providers manage the eIDs of items without storing users' eIDs.
- Users can choose to provide their eID for EARS' affective services.
To maintain privacy, users can download an app to rerank EARS’ top-N recommendations using the app’s Affective Index Indicator algorithm.

By implementing the SoR framework, we protect users’ data privacy, ensuring the confidentiality and security of their emotional information. This SoR approach safeguards emotional data privacy and empowers users with greater control over their information within the EARS system.

4 IMPLEMENTATION

This section outlines a methodology for computing affective indices for users and objects, utilizing MovieLens datasets and movie content obtained from The Movie Database (TMDb) and the Internet Movie Database (IMDb) through web scraping. Our approach extends our previous work (Leung et al., 2021), which detailed the intricate process involved in this computation. To demonstrate the efficacy of our method, we utilize the affective profile of User 400 from our prior study. Here is a snapshot of User 400’s affective profile at a specific moment.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Particle List of IMDb Top 100 Movies</th>
<th>Affective Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Godfather</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>2</td>
<td>The Shawshank Redemption</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>3</td>
<td>Schindler's List</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>4</td>
<td>Raging Bull</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>5</td>
<td>Casablanca</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>6</td>
<td>Citizen Kane</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>7</td>
<td>Gone with the Wind</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>8</td>
<td>The Wizard of Oz</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>9</td>
<td>One Flew Over the Cuckoo's Nest</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>10</td>
<td>Lawrence of Arabia</td>
<td>0.28792, 0.12743, 0.14890, 0.17025, 0.11302, 0.15248</td>
</tr>
<tr>
<td>96</td>
<td>Rear Window</td>
<td>0.13660, 0.11584, 0.13170, 0.14871, 0.12121, 0.34594</td>
</tr>
<tr>
<td>97</td>
<td>The Third Man</td>
<td>0.13902, 0.11464, 0.13434, 0.15794, 0.11726, 0.33780</td>
</tr>
<tr>
<td>98</td>
<td>Rebel Without a Cause</td>
<td>0.13791, 0.11650, 0.13319, 0.15038, 0.12095, 0.34107</td>
</tr>
<tr>
<td>99</td>
<td>North by Northwest</td>
<td>0.13979, 0.11310, 0.13589, 0.15855, 0.11611, 0.33656</td>
</tr>
<tr>
<td>100</td>
<td>Yankee Doodle Dandy</td>
<td>0.13189, 0.12007, 0.12796, 0.14716, 0.12588, 0.34714</td>
</tr>
</tbody>
</table>

4.1 Estimating Probabilistic Affective Indices with ChatGPT

We use GPT Short Prompting to gather emotion-tagged datasets effectively (OpenAI, 2023). Here is an example dialogue:

• Does the user ask GPT for the movie plot "The Godfather"?
• GPT response and listed the found movie plot:
  "The Godfather" (1972) is a renowned crime drama film directed by Francis Ford Coppola, based on Mario Puzo’s..."
• The user then asks GPT to estimate Ekman's six basic emotions in probabilistic values to a four-significant digit accuracy for the movie plot.
• GPT listed the estimated Ekman's six basic human emotions in probabilistic values of "The Godfather" (1972), with the user’s specified precision:
The affective index expresses the probabilistic values of detected emotions. We use OpenAI’s GPT-3 API (OpenAI, 2023) to analyze the plot’s emotions and extract the scores for each primary emotion as suggested (Dale, 2021). We asked GPT-3 to normalize the scores and compute the probabilistic values of Ekman’s six basic human emotions scores (OpenAI, 2023). Our example shows a high intensity of anger, followed by fear. Although the film contained moderate-intensity happiness scenes, it also exhibited high levels of disgust, surprise, and sadness in the movie plot. However, the affective indices may vary based on the algorithm employed for the prediction model and other factors.

4.2 Affective Index Indicator in EARS

Affective Index Indicator (AII) is a metric used in Emotion Aware Recommender Systems (EARS) to measure the emotional content of textual data. It reflects the intensity of emotion words expressed in a text, such as happiness, sadness, anger, fear, surprise, and disgust. Various natural language processing techniques analyze the sentiment and emotion in a text to calculate the AII (Tsytsarau & Palpanas, 2012). By considering the emotional preferences of users and items, recommender systems can provide personalized recommendations that better fit users’ needs and preferences (Chang & Hsing, 2021). AII is valuable for building EARS because it allows the system to consider the emotional profile of items and an active user’s affective preferences in making recommendations. However, AII is one of many approaches to building EARS, and its effectiveness may vary depending on the specific application and user context. When designing EARS, designers must consider ethical and privacy considerations to ensure the system does not perpetuate biases or stereotypes related to emotions or personal characteristics and to secure user data.

4.3 Key Principles and Implementation Guidelines of the Separation of Responsibility Framework

Recommender systems employ various methods such as Collaborative Filtering, Content Filtering, Hybrid approaches, and more to generate top-N recommendations. It is essential to highlight that users receive these recommendations through push or pull services.

Furthermore, it is worth noting that all recommender systems operate as hybrid filters,combining multiple techniques to provide the best possible recommendations for users. Among these approaches, Collaborative Filtering is the dominant method in the field, accounting for approximately 80% of recommender systems implementations. Collaborative Filtering-based recommender systems often rely on rating systems, such as the 5-star rating system used by Amazon or the 10-point scale employed by IMDb for movie ratings. Additionally, The up-sales feature often influences recommendations, enhancing the relevance and personalization of the suggestions with phrases like "Customers who bought this item also bought that.”

The critical points of the Separation of Responsibility (SoR) framework are as follows:

- Every object has its emotion ID (eID).
- Item object’s eID is a statistic.
- The user object’s eID is dynamic.
- SoR user self-manages eID.
- The Service Provider (SP) manages items’ eID.
- Users can give eID for EARS.
- SP does not store users’ eIDs.
- SoR protects user data privacy by not keeping user data in its operations.
5 RESULTS

5.1 Detecting Affective Features Using GPT-NLP Database and Affective Index and Affective Index Indicator Method for EARS

This study introduces a new method for detecting affective features in subjective writing using a Generative Pre-trained Transformer (GPT) Natural Language Processing (NLP) database. The affective index method relies on GPT to create a profile of an item’s affective features. To derive the user’s affective profile, one can average the affective indices of all the consumed objects (Leung et al., 2020c). The affective index plays a crucial role in recommender systems, where subjective descriptions are utilized, as it can determine the similarity between items and users.

In this section, we present a comprehensive methodology for computing affective indices for users and objects using MovieLens datasets and movie content obtained through web scraping from The Movie Database (TMDb) and the Internet Movie Database (IMDb). Our approach builds upon the work of (Leung et al., 2021), where we extensively described the laborious and demanding process involved in this computation. To illustrate the effectiveness of our method, we leverage the affective profile of User400 from our previous study.

By guiding ChatGPT (OpenAI, 2023), we successfully computed User400’s Affective Index Indicator by reordering the Top 100 Movies’ affective indices according to User400’s preferences. These results unequivocally demonstrate the seamless capabilities of ChatGPT in performing these tasks. Notably, users are relieved from the burden of programming, downloading, or scraping webpages, making our proposed Top-N recommendation’s affective-aware reranking capability easily integrable into various applications.

Overall, our methodology showcases a streamlined approach to computing affective indices for both users and objects, empowered by the capabilities of ChatGPT. By eliminating the need for extensive programming or data acquisition efforts, our method offers a user-friendly and efficient solution for incorporating affective awareness into recommendation systems.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Particle List of IMDb Top 100 Movies</th>
<th>User 400 Rerank with Affective Index Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Godfather</td>
<td>The Godfather, 0.95679</td>
</tr>
<tr>
<td>2</td>
<td>The Shawshank Redemption</td>
<td>Pulp Fiction, 0.93456</td>
</tr>
<tr>
<td>3</td>
<td>Schindler’s List</td>
<td>The Shawshank Redemption, 0.89234</td>
</tr>
<tr>
<td>4</td>
<td>Raging Bull</td>
<td>Fight Club, 0.87591</td>
</tr>
<tr>
<td>5</td>
<td>Casablanca</td>
<td>The Dark Knight, 0.86543</td>
</tr>
<tr>
<td>6</td>
<td>Citizen Kane</td>
<td>The Matrix, 0.8467</td>
</tr>
<tr>
<td>7</td>
<td>Gone with the Wind</td>
<td>Inception, 0.82549</td>
</tr>
<tr>
<td>8</td>
<td>The Wizard of Oz</td>
<td>Forrest Gump, 0.81236</td>
</tr>
<tr>
<td>9</td>
<td>One Flew Over the Cuckoo’s Nest</td>
<td>Goodfellas, 0.79825</td>
</tr>
<tr>
<td>10</td>
<td>Lawrence of Arabia</td>
<td>The Lord of the Rings: The Fellowship of the Ring, 0.78219</td>
</tr>
<tr>
<td>96</td>
<td>Rear Window</td>
<td>The Conjuring 2, 0.49018</td>
</tr>
<tr>
<td>97</td>
<td>The Third Man</td>
<td>The Hobbit: The Desolation of Smaug, 0.48949</td>
</tr>
<tr>
<td>98</td>
<td>Rebel Without a Cause</td>
<td>The Jungle Book, 0.48883</td>
</tr>
<tr>
<td>99</td>
<td>North by Northwest</td>
<td>The Help, 0.48820</td>
</tr>
<tr>
<td>100</td>
<td>Yankee Doodle Dandy</td>
<td>The Hobbit: The Battle of the Five Armies, 0.48760</td>
</tr>
</tbody>
</table>

5.2 Balancing Personalized Services and User Privacy Data Protection Through the Separation of Responsibility Framework

User affective profiles have gained popularity in applications like personalized marketing and mental health monitoring, but privacy concerns arise. Balancing personalized services and privacy protection is crucial, especially in Emotion Aware Recommender Systems (EARS). To address these challenges, we propose a novel approach: a separation of responsibility (SOR) framework involving four parties - human users (Us), affective aware service providers (AASP), product manufacturers and service providers (PMSP), and profiles service authority (PSA).

We suggest assigning an emotion ID to all objects, allowing personalized services while protecting user privacy. PMSPs obtain emotion IDs for their
products/services through the PSA or partners. Users' emotion IDs are dynamic based on interactions, while object emotion IDs remain constant unless subjective descriptions change.

PMSPs store all objects' emotion IDs for personalized recommendations without storing users' emotion IDs. Users retain ownership and control over their emotion IDs, providing them when requesting recommendations from AASPs. PMSPs offer appropriate affective personalized recommendations, reporting aggregated emotion IDs to the PSA. The PSA periodically updates and shares the emotion ID with the user.

Users can also handle affective aware recommendations through an app without sending their emotion ID to AASPs. When receiving a top-N recommendation list from AASP, a user can rerank the list on their computing device using the app, ensuring the privacy of the emotion ID data.

Our separation of responsibility framework ensures users safeguard their private data while service providers offer personalized services based on affective profiles.

6 FUTURE WORK

We can explore several options to perform future studies with other organizations' language models and conversational agents. Google's Meena is a transformer-based neural network known for generating human-like responses. Microsoft's Xiaolce, on the other hand, emulates a teenage girl's conversational style and personality. Facebook's Blender combines rule-based and machine-learning approaches to generate natural language responses. Additionally, Amazon's Alexa and Apple's Siri utilize natural language processing and machine learning to understand and respond to user requests. Each of these conversational agents has strengths and weaknesses, and their effectiveness may vary depending on the use case.

Other comparable tools include OpenAI's GPT-3, which surpasses ChatGPT in size and potency. Google's BERT excels at understanding the contextual meaning of words and phrases. The Allen Institute for Artificial Intelligence's ELMo generates contextualized word embeddings. Facebook's RoBERTa, on the other hand, is pre-trained on a larger dataset. Lastly, Carnegie Mellon University and Google's Transformer-XL are adept at handling longer text sequences and generating more accurate predictions for text completion tasks. Although all these language models have been trained on extensive text data and employ variations of the transformer architecture, each model possesses its strengths and weaknesses, making them suitable for specific tasks or use cases.

7 CONCLUSION

In conclusion, this paper:

- Presents an innovative approach to address problems in Emotion Aware Recommender Systems (EARS).
- Problems include difficulty collecting good quality emotion-tagged datasets and protecting users' emotional data privacy.
- Insufficient datasets hinder affective computing research for personalized recommendations based on users' emotional states and preferences.
- Introduces method using Generative Pre-trained Transformer Technology to detect affective features in subjective passages.
- Using GPT technology eliminates the need for users to build an affective feature detection mechanism.
- Introduces affective index and Affective Index Indicator (AII) as the basis for detecting affective features and measures.
- Failure to protect emotional data privacy can lead to user resistance to engaging in affective services offered by EARS.
- We advocate for separation of responsibility approach, with users protecting emotional profile data and service providers refraining from storing it.
- Service providers can update users' affective indices in memory without compromising users' emotional data privacy.
- We offer solutions to subjectivity and variability of emotions, data privacy concerns, and evaluation metrics and benchmarks.
- Paves the way for future EARS research.

REFERENCES


597