A Personalized Hybrid Tourist Destination Recommendation System: An Integration of Emotion and Sentiment Approach

Suphitcha Chanrueang[1], Sotarat Thammaboosadee[2]*, Hongnian Yu[3]
Faculty of Engineering, Mahidol University, Nakhon Pathom, Thailand[1, 2]
School of Computing Engineering and the Built Environment, Edinburgh Napier University, Edinburgh, United Kingdom[3]

***Abstract*—This research introduces a personalized hybrid tourist destination recommendation system tailored for the growing trend of independent travel, which leverages social media data for trip planning. The system sets itself apart from traditional models by incorporating both emotional and sentiment data from social platforms to create customized travel experiences. The proposed approach utilizes Machine Learning techniques to improve recommendation accuracy, employing Collaborative Filtering for emotional pattern recognition and Content-based Filtering for sentiment-driven destination analysis. This integration results in a sophisticated weighted hybrid model that effectively balances the strengths of both filtering techniques. Empirical evaluations produced RMSE, MAE, and MSE scores of 0.301, 0.317, and 0.311, respectively, indicating the system's superior performance in predicting user preferences and interpreting emotional data. These findings highlight a significant advancement over previous recommendation systems, demonstrating how the integration of emotional and sentiment analysis can not only improve accuracy but also enhance user satisfaction by providing more personalized and contextually relevant travel suggestions. Furthermore, this study underscores the broader implications of such analysis in various industries, opening new avenues for future research and practical implementation in fields where personalized recommendations are crucial for enhancing user experience and engagement.**

***Keywords*—Recommendations; hybrid recommendation system; Collaborative Filtering; Content-based Filtering; social media data; travel planning**

I. Introduction

Tourism significantly boosts the economy, creates jobs, and reduces poverty through spending, investments, and government backing [1]. Effective government policies are crucial for maximizing tourism's benefits and fostering overall economic growth. Social media , especially Facebook, plays a pivotal role in influencing travel choices and providing essential information to travelers [2]. Many travelers primarily rely on blogs and vlogs [3] for travel inspiration.

Emotions expressed on Facebook, such as anger, sadness, fear, joy, and love, can be identified through advanced textual analysis techniques [4]. Among these, anger, sadness, and fear are the most common, while love and amazement top the reactions [5]. Sentiment analysis, powered by sophisticated machine learning algorithms, evaluates sentiments on Facebook by analyzing data from the Top Page [6]. Understanding user perceptions and engagement is crucial for informed decision-making on social media platforms. This technique can also be applied to other digital platforms to assess customer feedback sentiment. Automating this process facilitates the examination of large datasets, effectively addressing the inherent challenges in sentiment analysis [7].

Recommender systems are instrumental in helping tourists discover attractions that match their personal preferences and needs. These systems utilize various techniques, including social and Bayesian networks, Collaborative Filtering (CF) algorithms, and deep learning models, to analyze user behavior, textual sentiment, and similarities among options. By considering user characteristics, behaviors, social network connections, and search contexts, these systems can make precise recommendations for tourist destinations that align with individual interests [8].

In this study, we introduce a personalized hybrid tourist destination recommendation system that leverages emotional and sentiment data from social media platforms. Unlike conventional models, our system integrates these emotional cues to provide more nuanced and accurate travel suggestions. Our system combines CF and Content-based Filtering (CB) with sentiment analysis techniques. CF is used to recognize emotional patterns, while CB analyzes sentiment-driven data, resulting [9] in a sophisticated weighted hybrid model [10]. This approach ensures recommendations are finely tuned to capture the nuanced emotional responses of users.

By integrating emotional and sentiment analysis, our system enhances user satisfaction by adapting to changes in user preferences over time, providing dynamic and contextually aware recommendations. This leads to a more engaging and satisfying user experience. The key advancements of our model include dynamic weighting of user data, enhanced emotional resonance, and adaptability. These improvements make our model more accurate and personalized compared to traditional methods.

Overall, the sentiment and emotion-based Weighted Hybrid technique not only improves the technical robustness of recommender systems but also significantly elevates their practical application by delivering a more personalized, accurate, and emotionally resonant travel experience. This advancement represents a substantial leap forward in the field of tourism recommendation systems, setting a new standard for personalized travel planning.

In terms of technical aspects, the proposed model combines CF and CB with sentiment analysis techniques. CF is used to recognize emotional patterns, while CB analyzes sentiment-driven data, resulting in a sophisticated weighted hybrid model. This approach ensures recommendations are finely tuned to capture the nuanced emotional responses of users. To the best of our knowledge, no previous research has utilized emotions derived from social media reactions in their recommendation systems.

Regarding practical applications, by integrating emotional and sentiment analysis, the developed approach enhances user satisfaction by adapting to changes in user preferences over time, providing dynamic and contextually aware recommendations. This leads to a more engaging and satisfying user experience. The key advancements of this model include dynamic weighting of user data, enhanced emotional resonance, and adaptability. These improvements make the proposed model more accurate and personalized compared to traditional methods.

Therefore, the objectives of the presented paper are to:

1) Extract and analyze data focusing on the emotions and sentiments in posts, comments, and reactions about tourist spots to discern their influence on travel decisions.

2) Develop a state-of-the-art hybrid recommender system that combines CB and CF with sentiment analysis from Facebook data, offering personalized suggestions for tourist destinations.

The paper is structured into six sections, including a comprehensive review of related works in Section II, system design and methods in Section III, experimental results in Section IV, a discussion of these findings in Section V, and conclusions and future directions in Section VI.

II. Related Works

A. Recommender Systems

Recommendation systems have become indispensable across numerous sectors, including e-commerce, entertainment, news, and social networking, by facilitating access to tailored information and resources. These systems streamline the search process, allowing users to find resources suitable to their needs by providing individualized suggestions or guiding them to relevant resources within a large data space. They simplify the process of finding information and solutions, making it easier for customers and project providers to identify and receive projects and other services [11]. In the tourism sector, these systems assist users in locating resources that match their specific requirements by offering personalized recommendations or directing them towards pertinent resources within a vast data environment [12]. By analyzing user preferences and behavior, they filter and present tailored options, significantly reducing the time and effort needed to find relevant information or items in an otherwise overwhelming data landscape [13]. They also play a crucial role in guiding customers throughout their shopping journey, presenting the most relevant products without the need for explicit searches [14]. By incorporating ontologies and machine-learning algorithms, recommender systems enhance accuracy and efficiency [15], addressing challenges and improving business productivity.

Recommendation systems can be categorized into several types, each with a unique approach to providing personalized recommendations. CF is one of the most common methods [16] and can be divided into user-based and item-based approaches. User-based Collaborative Filtering recommends items based on the preferences of users who have similar tastes, while Item-based Collaborative Filtering suggests items that are similar to those the user has previously liked or interacted with. CB focuses on recommending items that share similar attributes or features with those the user has shown interest in. Hybrid Systems (HS) combine multiple recommendation techniques, such as CB and CF, to enhance the overall accuracy and relevance of the recommendations. Context-aware Recommender Systems consider contextual factors such as time, location, or current activity to tailor recommendations more closely to the user's present situation. Demographic Recommender Systems provide suggestions based on demographic data, such as age, gender, or education level.

The research methodologies employed in tourism recommendation systems exhibit considerable diversity. Some studies focus on analyzing and quantifying user sentiment toward tourism destinations based on text reviews [17], integrating these sentiments [18] into the recommendation model. Others employ hybrid methods that combine CB and CF, utilizing preprocessed data from websites for recommendation. Additional approaches include probabilistic topic modeling and custom day itinerary models to analyze tourist travel patterns and preferences. While some studies emphasize recommending points of interest within a tourist attraction based on visitor interests, others offer broader recommendations spanning entire countries.

In summary, personalized hybrid recommendation systems across various domains leverage individualized suggestions and advanced techniques like opinion mining and hybrid filtering (HF) to enhance accuracy [19] and user experience. Despite their effectiveness in simplifying information discovery and improving the customer journey, these systems encounter challenges related to relevance computation, personalization, and the integration of specific user interests within large data spaces.

B. Factors Influencing Personal Travel Destination Choices

The decision-making process regarding travel destinations is influenced by a complex interplay of factors that vary significantly among individuals based on their preferences and circumstances, timing of the travel, and the quality of infrastructure and traffic conditions, which collectively shape the feasibility [20] and appeal of a destination [21]. The broad availability of information from different channels like online platforms, print media, and travel agencies plays a crucial role in informing potential travelers about their options, thereby significantly influencing their destination choices [22]. Moreover, the operational efficiency and overall attractiveness of a tourism destination, determined by factors like labor quality, capital investment, technological advancement, environmental sustainability, financial expenditure, generated revenue, and the potential length of stay, are critical in swaying personal travel destination choices [23]. These considerations encompass a range of practical, economic, and subjective factors that contribute to the appeal and competitiveness of a destination, highlighting the multifaceted nature of travel decision-making. In essence, the choice of a travel destination emerges from a dynamic balance of these practical considerations, individual preferences, and the intrinsic attributes of the destination itself, underscoring the complexity of travel planning and the importance of understanding these factors for stakeholders in the tourism industry.

C. *Sentiment and Emotion as New Factors for Recommendation Systems in the Tourism Domain*

Sentiment and emotion play a crucial role in enhancing recommendation systems [24], [25] in the tourism domain. Incorporating sentiment analysis from user-generated content like reviews can significantly improve the accuracy [26] and quality of recommendations. By utilizing sentiment and emotion scores derived from user reviews, tourism recommendation systems can better capture user preferences and generate personalized recommendation lists based on semantic similarity [27]. Additionally, aspect-based sentiment analysis models can extract sentiment polarity from reviews, providing insights into tourists' evaluations and enhancing service and product upgrades [28].

These sentiment-driven approaches not only help in understanding tourists' emotions but also assist tourism organizations in their decision-making processes, ultimately leading to more effective recommendations [29] and greater customer satisfaction.

D. *Weight Hybrid Recommendation System*

A weighted hybrid model in recommendation systems combines multiple techniques, such as CF and CB, by assigning different weights to each technique based on their effectiveness in predicting user preferences. This approach leverages the strengths of each method to enhance recommendation accuracy and relevance [30]. The model integrates CF, which identifies patterns based on user interactions, and CB, which recommends items with similar attributes to those the user likes. Each method generates a recommendation score, and their contributions are weighted differently, with the weights determined [31] through experiments or data characteristics. One of the key advantages of a weighted hybrid model is its flexibility. The weights can be adjusted dynamically based on the recommendation context, user behavior, or data changes. This adaptability helps improve recommendation relevance over time, addressing limitations like the cold start problem and data sparsity [32], and providing more accurate and diverse suggestions.

In summary, a weighted hybrid model strategically combines multiple recommendation techniques with assigned weights to enhance accuracy and personalization. This method leverages the strengths of different techniques, adapts to changing data and user behaviors, and provides a robust and personalized recommendation experience.

III. SYSTEM DESIGN AND METHODS

The model aims to develop an innovative personalized tourist attraction recommendation system, showcasing a novel architecture that incorporates advanced HF techniques, as illustrated in Fig. 1. This system leverages three distinct types of information, marking a significant advancement in tourist recommendation technologies.

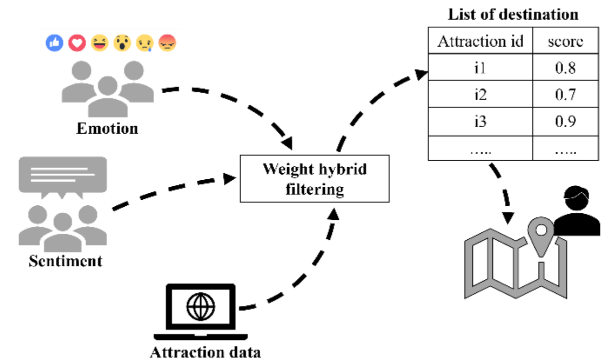

Fig. 1. Proposed architecture for a tourism recommender system.

A. *Data Collection and Authorization*

In collaboration with five Thai Facebook fan pages, we secured authorization to extract a wealth of data, including comments from followers, their emotional reactions via the 'reaction' button, and detailed information on various tourist attractions. This comprehensive dataset, accumulated two years, offering deep insights into tourist preferences and behaviors.

The architecture of a Personalized Hybrid Tourist Destination Recommendation System, as depicted in Fig. 1, employs a hybrid approach that integrates both CB and CF models. An in-depth explanation of the main segments and their respective processes is provided in multiple sections, as shown in Fig. 2.

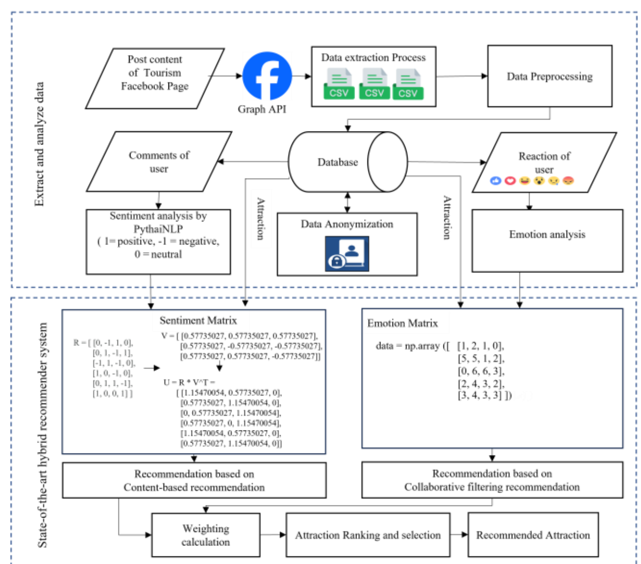

Fig. 2. Architecture of a weighted hybrid recommendation system.

B. *Data Collection Process*

As illustrated in Fig. 2, the architecture of a WHF begins with the data collection process, which gathers information from various sources. One such source is Facebook, where data is collected using the Graph API. This data, which ranges from user demographics to their interactions with content, is typically stored in a CSV format to facilitate handling and analysis.

The initial stage of data preprocessing involves extracting and refining the gathered information through a series of steps. This phase focuses on ensuring data quality by cleaning inaccuracies, making necessary corrections, and identifying essential information to enable accurate recommendations. In respect of user privacy, data anonymization is implemented. During this process, data is stripped of personal identifiers, ensuring user privacy protection while still allowing for personalized content suggestions.

The system performs sentiment analysis by categorizing user comments as positive, neutral, or negative. Additionally, it evaluates user reactions, such as likes or emojis, by assigning numerical values from one to six to quantify user engagement. The processed data is organized into a matrix format, where users are listed alongside the items they interact with, creating a comprehensive map of interactions.

C. Data Selection Criteria

Selecting data for a tourism recommendation system is a complex endeavor due to the sheer number of tourism attractions and the overwhelming volume of information available online and across social media platforms. Existing recommender systems encounter challenges in delivering precise recommendations, as they must contend with variations in users' interests, the ever-changing contexts, and the sequential patterns of travel [33]. The lack of sufficient historical user data in the tourism sector further complicates matters, leading to difficulties such as cold starts and data sparsity, which hinder the delivery of accurate and reliable recommendations [34]. Furthermore, the infrequent browsing and purchasing of travel products, along with the influence of factors such as departure, destination, and price, adds another layer of complexity to recommending travel products [35].

Our method for selecting sentiment and emotion data follows strict criteria to ensure its relevance, accuracy, and diversity, while maintaining privacy and ethical standards. We identify key emotional data like user comments and reactions, verify their accuracy, and source them from various platforms, including social media. This data must be scalable and comprehensive to support reliable analysis and improve the system's ability to offer accurate and trustworthy recommendations, thereby enhancing its effectiveness.

D. Data Extraction Process

In social media data extraction, using Facebook's Graph API is crucial for researchers and technologists retrieving data from fan pages. Facebook Graph API allows developers to access and interact with Facebook data, such as user profiles, posts, and photos, using HTTP requests. It requires authentication via access tokens for secure data access and supports CRUD operations. This API enables the integration of Facebook data into applications for social media management, analytics, and personalized content delivery. Our study utilizes Facebook's Graph API [36] as a key tool, following a systematic method that values user privacy and adheres to privacy regulations , as shown in Fig. 3.

Fig. 3. Data preparation process.

The process starts by configuring an application on Facebook Developer Console to obtain an App ID and App Secret, enabling user consent through OAuth 2.0 for an access token. The access token allows fetching data from fan pages using Graph API, which is crucial for analysis.

Data retrieval involves accessing posts, reactions, and comments from fan pages through Graph API to gather raw data for analysis and recommendations. Anonymization techniques are applied to protect user data, including removing identifiers, randomizing sensitive data, and auditing the process regularly. The system securely stores anonymized reactions and comments in a privacy-compliant database and manages them through an ETL pipeline to maintain anonymization. Utilizing Facebook's Graph API involves access setup, data retrieval, and strict anonymization, laying the groundwork for further data processing in the Data Preparation phase.

E. Data Preparation

The data preparation phase involves handling three types of raw data crucial for constructing the dataset.

1) Social network data: Provides attraction names from social networks to identify and categorize tourist destinations.

2) Sentiment data transformation: Categorizes opinions from user reviews into three sentiment categories using the PyThaiNLP library [38] for sentiment analysis tailored to the Thai language. This library is vital for tasks specific to Thai, such as word tokenization for interpreting user sentiments accurately.

3) Model for identifying and scoring user emotions: This model identifies core emotions such as anger, disgust, fear, happiness, sadness, and surprise. The proposed model is consistent with Ekman's framework and is widely acknowledged in the field of emotion recognition. Emotions are scored to reflect user engagement [39]: Love = 6, Like = 5, Haha = 3, Wow = 4, Sad = 2, and Angry = 1, providing insights into the subtleties of user emotional feedback. The process is illustrated in Fig. 4.

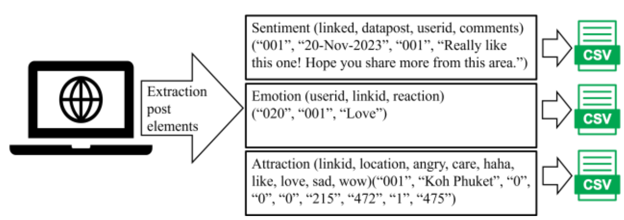

Fig. 4. Data preparation process.

Social network and user review data hold implicit details, requiring specific feature extraction techniques. Social computing retrieves social information from social network data, whereas sentiment analysis uncovers emotional cues from user feedback. Data collection and integration procedures prioritize safeguarding user privacy by anonymizing sensitive information. These privacy steps are vital for upholding trust and efficiency in our recommendation system, offering personalized suggestions while protecting user privacy.

F. *Creation of Recommendation System*

The creation of the recommendation system involves the integration of CF and CB models. Below, we detail the methodologies used for each model:

1) Collaborative Filtering model: The development of a tourist attraction recommendation system using CF with the Singular Value Decomposition (SVD) algorithm is achieved through the following steps, with conceptual underpinnings and practical implementation in Python. Installation and Setup: Begin by installing the Surprise package using a package manager like pip, and import necessary classes such as Dataset, Reader, SVD, and accuracy functions from the library.

Conceptual Framework of SVD: For a matrix A with dimensions $m \times n$, SVD decomposes A into three matrices (1):

$$R = U \cdot \Sigma \cdot V^T \tag{1}$$

U: User ser features matrix, where rows represent users and columns represent hidden characteristics.

Σ: Diagonal matrix of singular values indicating the importance of the latent features.

V^T: Item-feature matrix, with row for items (attractions) and columns for latent features.

The procedure includes constructing matrices for User factors and Item factors, as illustrated in Fig. 5. Based on the example data provided, this step entails training a model using SVD.

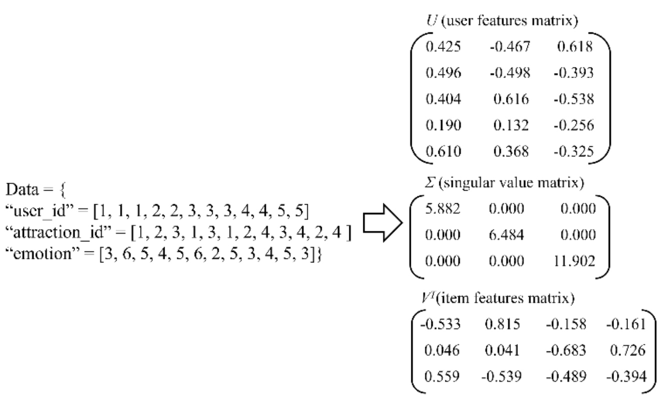

Fig. 5. Matrices are generated for user factors and item factors.

Data Preparation: Prepare the data, including 'User_id', 'Attraction_id', and 'emotion' scores ranging from 1 to 6. An example data structure is provided.

Model Training: Define the range of 'emotion' scores using the Reader class, setting the minimum and maximum values. Load the data using the Dataset module, formatted according to the Reader specifications. Split the data into a training set for training the model and a test set for evaluating its performance. Instantiate the SVD algorithm and fit it to the training dataset.

Prediction and Evaluation: The system makes predictions for unseen user-attraction combinations in the test set. For example, it predicts the 'emotion' score for a given user-attraction pair, such as 'User123' (user_ID) and 'Attraction456' (attraction_ID). If the actual emotion score given by the user is 4 and the score estimated by the SVD model is 3.8, this indicates the model's performance. The details section, which shows 'was impossible': False, confirms that the prediction was successfully computed. To evaluate the model's accuracy, metrics such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are used [40]. These metrics compare the predicted scores against the actual scores in the test set. RMSE measures the square root of the average squared differences between the predicted and actual values, while MAE measures the average of the absolute differences between the predicted and actual values.

Practical Considerations: To protect user privacy, data anonymization is crucial. This involves removing or obscuring personal identifiers from the data to prevent individual users from being easily identified. Regular audits of the anonymization process are essential to ensure the data remains secure and to minimize the risk of re-identification, maintaining user trust and compliance with privacy regulations.

2) Content-based filtering model: Implementing a CB model for travel recommendations with Support Vector Machine (SVM) involves a structured approach, focusing on harnessing powerful capabilities in handling complex data patterns. The key steps in this process include data preparation, model selection and tuning, training, and evaluation.

Data Preparation: Transform sentiment data into numerical scores of (-1, 0, 1) to align with the SVM models' requirements. This step converts subjective sentiments into objective data points. Convert characteristics of tourist attractions, such as type and location, into numerical forms labeled as 'Attraction_type'. This numerical transformation is crucial for machine learning algorithms to process and learn from the data.

Model Selection and Tuning: Choosing and tuning the model is crucial as we opt for SVM due to its proficiency in classification tasks, especially its effectiveness in high-dimensional spaces and its ability to handle non-linear data separation through kernel methods. Fine-tune parameters like C (regularization), `kernel type`, and `gamma` (kernel coefficient) to optimize the model for the dataset. In classification tasks, the primary objective is to find an optimal hyperplane that best separates the classes in the given dataset.

The SVM algorithm seeks to find an optimal hyperplane that best separates different classes in the dataset, as expressed by the hyperplane Eq. (2):

$$w^T x + b = 0 \tag{2}$$

where w represents the weight vector, x is the vector of data points, and b signifies the bias. Alongside the hyperplane equation, SVM involves an optimization problem, which is geared towards maximizing the margin between the data classes, as shown in Eq. (3):

$$min(w, b) \left(\frac{1}{2} ||w||^2\right) \tag{3}$$

Subject to the constraints for each data point *i*:

$$y_i(w^T x_i + b) \geq 1, \forall i \quad (4)$$

In this context: $||w||2$ is the norm of the weight vector, and minimizing it is key to maximizing the margin. The labels of the data points are denoted by y_i and x_i represents each data point.

Training and Evaluation: During model training, we divide our dataset into training and test sets to both train the model and evaluate its predictive performance accurately. After training, we assess the model's performance using metrics such as accuracy, precision, recall, and F1-score. These metrics provide a comprehensive understanding of the model's effectiveness in classification tasks.

Implementation: Once the model proves its effectiveness, it can be utilized to forecast user preferences for different tourist attractions and provide suitable recommendations. Utilizing SVM enables us to adjust and explore the model for optimal performance on our dataset. This methodology is especially beneficial for tasks demanding nuanced data analysis.

Visualize the sentiment feature matrix in SVM as shown in Table I to understand the data's distribution and how the model determines decision boundaries. Once the model is optimized and validated, it can accurately predict user preferences for various tourist attractions, providing personalized recommendations based on these analyzed features.

1) Weight hybrid recommendation: Integrating CB with CF in a WHS involves a structured approach to leverage the advantages of both methods. The workflow is as follows*:*

TABLE I. SENTIMENT FEATURE MATRIX OF SVM

Seq	Attraction_id	User_id	Sentiment
0	101	201	-1
1	102	202	0
3	104	204	-1
4	105	201	-1
5	105	202	0

a) Data processing and score calculation: Initially, both CB and CF systems process their respective datasets to compute scores for tourist destinations. These scores are based on each system's unique algorithms and the data provided.

b) Blending scores: The critical step of blending involves merging the scores from both systems using a predefined formula. This formula assigns specific weights to the scores from each system, balancing their contributions. For instance, the hybrid score can be calculated as:

$$\Sigma\chi o\rho\varepsilon_{\mathrm{H}\psi\beta\rho\iota\delta} = \alpha \times \Sigma\chi o\rho\varepsilon_{\mathrm{XB}} + (1-\alpha) \times \Sigma\chi o\rho\varepsilon_{\mathrm{X}o\lambda\lambda\alpha\beta o\rho\alpha\tau\iota\varpi\varepsilon} \quad (5)$$

Here, α represents the weight assigned to the score from the CB Filtering system.

1) Optimizing the weight parameter (α): Fine-tuning α is crucial for balancing CB and CF systems. Experimenting with various α values, calculating blended scores, and analyzing outcomes enhances recommendation accuracy and diversity. K-grid tuning optimizes model parameters by adjusting the K value for cross-validation groups. Selecting an optimal K value, evaluating model performance, and refining based on results analysis ensures models align with data characteristics and task requirements.

2) Performance evaluation: After the hybrid recommendation system is in place, its performance should be evaluated to ensure it provides relevant and accurate suggestions. Utilizing feedback from users and analyzing performance metrics such as Accuracy, Precision, Recall, and F1-Score are integral to this phase. These evaluations facilitate ongoing refinement, improving the system's capability to effectively cater to user preferences.

This methodology underscores the importance of a balanced integration of CB and CF techniques, ensuring that the recommendations are not only accurate but also varied, catering to the diverse interests of users.

IV. EXPERIMENTAL RESULTS

The analysis of recommendation results derived from a comprehensive dataset provides valuable insights for the development and evaluation of a tourist attraction recommendation system. The following discussion presents a structured analysis based on the dataset results.

A. Dataset Results

The data extraction process leveraged a fan page dedicated to tourist attraction reviews, a valuable resource for assessing recommendation models. The dataset's composition and its implications for the recommendation system are as follows:

1) Dataset overview: The dataset contains 252,568 records split into two segments: 151,541 records for training and 101,027 for testing. This division ensures a robust framework for both developing and validating the recommendation model.

2) User participation and attraction diversity: A total of 38,739 users have contributed to the dataset, reviewing 508 different attractions. Examples of the data can be found in Tables II to IV. This level of participation and variety underscores the dataset's richness and diversity, providing a solid foundation for generating nuanced and wide-ranging recommendations.

3) Sentiment classification: Sentiments extracted from the dataset are categorized into three categories: positive, negative, and neutral, as indicated in Table II. This classification facilitates a detailed understanding of user preferences and emotions regarding various attractions.

TABLE II. EXAMPLE OF SENTIMENT DATA

User Id	Attraction		Sentiment	
	Id	Type	Type	Score
1	157	1	Neutral	0.5
1	122	5	Positive	0.8
2	28	1	Positive	0.7
3	89	6	Positive	0.6
3	210	2	Neutral	0.5

4) Emotion ratings: User ratings are based on an emotional scale from 1 to 6, where each number corresponds to a specific emotion (e.g., love = 6, like = 5, wow = 4, haha = 3, sad = 2 and angry = 1), as illustrated in Table III.

TABLE III. EXAMPLE OF USER EMOTION DATA

User ID	Attraction		Emotion	
	Id	Type	Word	Value
1	627	2	wow	4
1	781	5	like	5
3	783	6	wow	4
3	210	2	like	5
7	157	1	love	6

The dataset shown in Table IV includes a variety of attraction types, illustrating the diverse interests of the users. Understanding the range of attractions is crucial for tailoring recommendations to suit individual user preferences effectively.

The following table, Table V, lists different types of attractions along with their corresponding type names.

TABLE IV. EXAMPLE OF ATTRACTION DATA

Attraction ID	Name		Location	Type
1	Khun Dan Prakarn Chon Dam		14.7994° N, 98.5969° E	4
2	Sai Yok National Park		14.417778° N, 98.747222°E	4
3	Singha Historical Park		14.03583° N, 99.23972°E	3
4	Phra Pathom Chedi		13° 49' 6.59" N, 100° 03' 22.20" E	6

TABLE V. EXAMPLE OF ATTRACTION TYPE DATA

Attraction Type	Type Name
1	Eco tourism
2	Arts and sciences educational attraction
3	Historical attraction
4	Natural attraction
5	Recreational attraction
6	Cultural attraction

B. Recommendation Results

The recommendation results in Table VI provide a comparative analysis of the performance of three models: CF, CB, and HF using Precision, Recall, and F1-score metrics. The HF model outperforms both CF and CB in all three metrics. This superior performance can be attributed to its ability to combine the strengths of both CF and CB techniques, along with additional enhancements such as sentiment and emotion analysis. This integration allows the HF model to provide more accurate and reliable recommendations, better predicting user preferences and enhancing the overall user experience in tourism settings. In conclusion, the HF model is the most effective approach for recommending tourist attractions, as it achieves the highest precision, recall, and F1-score, significantly improving recommendation accuracy compared to the CF and CB models.

TABLE VI. RESULT OF PRECISION, RECALL AND F1- SCORE

Model	Precision	Recall	F1-score
CF	0.780	0.740	0.760
CB	0.660	0.690	0.670
HF	0.850	0.830	0.840

This analysis compares the efficacy of three recommendation models: CF, CB Filtering, and HF, using Precision, Recall, and F1-score as performance metrics. Fig. 6 displays the trends of these metrics as the value of K changes, illustrating how different parameter settings affect model performance.

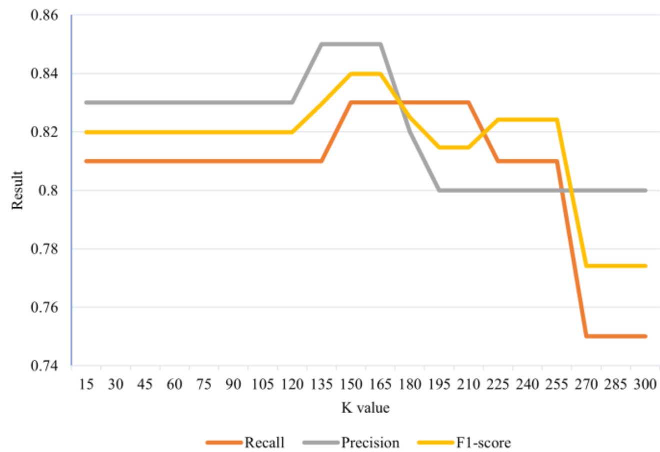

Fig. 6. Model discrimination score by K-value.

Content-Based Filtering vs. Collaborative Filtering: CF demonstrates robust performance with a well-balanced trade-off between Precision (0.78) and Recall (0.74), resulting in an F1-score of 0.76. In contrast, CB Filtering, while slightly less effective, demonstrated modest performance with Precision (0.66) and Recall (0.69), and an F1-score of 0.67. This suggests a modest decline in performance compared to CF. The HF model, which combines CB and CF, demonstrated superior performance across all metrics. It achieved the highest Precision (0.85), Recall (0.83), and F1-score (0.84), signifying its effectiveness in providing accurate and comprehensive recommendations.

The accuracy of a recommendation system, which measures how precisely it predicts user preferences, can be assessed using metrics like the RMSE or MAE [37]. These metrics, along with the MSE, provide insight into the system's performance and can be calculated using the functionalities available in Scikit-learn [41]. The results of these calculations are presented in Fig. 7.

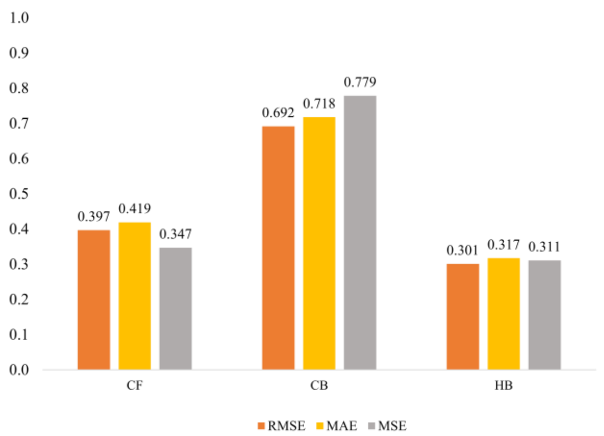

Fig. 7. Model performance comparison.

Error Rate Analysis: The analysis of RMSE, MAE, and MSE metrics presented in Fig. 7 revealed that the CB model exhibited the highest error rates, suggesting limitations in its predictive accuracy. The CF model showed moderate error values, indicating a balanced yet not optimal level of accuracy. The Hybrid model outperformed both with the lowest error values in all metrics, indicating its superior accuracy in predicting user preferences. This summary indicates that the HF model, which combines elements of both Collaborative and CB methodologies, offers the most effective approach for accurate user preference prediction.

C. *Comparison of Methods, Items, and Values Results*

The proposed method demonstrates the highest efficiency in recommendation systems, exhibiting the lowest error rates across all evaluated metrics, including RMSE, MAE, and MSE. This method integrates sentiment and emotion analysis with CF and CB techniques, leading to a significant improvement in recommendation accuracy. The integration of these emotional and sentimental data allows the system to better understand and predict user preferences, resulting in more personalized and precise recommendations.

In comparison, other methods such as those referenced in, [21], [30], and [42], utilize various combinations of CF and CB techniques, sometimes incorporating additional methods like SVD and SVM. Despite these efforts, they still exhibit higher error rates. For instance, the method in [21], which uses a hybrid approach with SVD and weighted techniques, shows better performance than some but still falls short compared to the proposed method. The methods in [33] and [42], although incorporating diverse techniques, demonstrate even higher error rates, indicating less accuracy in their recommendations. All the data is detailed in Table VII.

Overall, the proposed method's ability to integrate emotional and sentiment analysis with traditional filtering techniques sets it apart, achieving superior performance and underscoring the importance of these factors in enhancing recommendation systems.

TABLE VII. COMPARISONS WITH RECENT METHODS

Ref	Method	Technique	Result		
			RMSE	**MAE**	**MSE**
[21]	Hybrid	SVD Weighted	0.500	0.414	0.254
[30]	CB+CF	Weighted	0.880	0.670	-
[42]	Hybrid	Cosine, SVD, SVM	0.864	0.666	-
Proposed Method	Baseline+ CB+CF	SVD, SVM, Sentiment, Emotion	0.301	0.317	0.311

V. DISCUSSION

This study demonstrates the potential of a Hybrid Filtering method for enhancing efficiency and accuracy in tourist attraction recommendations. By intelligently combining CB and CF techniques with adjustable method weights, the proposed approach delivers highly personalized recommendations that align closely with individual user preferences. Previous research has highlighted the strengths of CB and CF methods individually, but the integration of these techniques with customizable weights offers a novel approach that addresses limitations in prior studies [8], [10]. Despite these advancements, the methodology encounters significant challenges in sentiment analysis and data extraction from social media platforms, particularly Facebook.

A. *Sentiment Analysis Challenges*

The study acknowledges the inherent complexities in interpreting emotions expressed on social media, consistent with findings from existing literature [4], [6]. Social media users often present idealized versions of their emotions, which may not accurately reflect their true sentiments [20]. Additionally, the diversity of content types (text, images, videos) and nuanced language used on these platforms further complicate sentiment interpretation [27]. These challenges underscore the need for advanced sentiment analysis tools capable of understanding diverse expressions and cultural contexts [26]. Future research could build on recent advancements in sentiment analysis techniques to improve interpretation accuracy [18].

B. *Data Extraction Complexities*

Relying on Facebook's Graph API for data retrieval introduces significant challenges, a problem well-documented in the literature [4]. Researchers must navigate strict personal data access restrictions, frequent API changes, data request limits, and complex verification processes, all while managing privacy risks. The complexity of ensuring compliance with Facebook's policies adds another layer of difficulty, requiring careful data transformation to protect personal information—a process that is often time-consuming. While the WHF method demonstrates superior performance by effectively leveraging data from multiple sources, it still faces substantial hurdles in sentiment analysis and social media data extraction. These findings are consistent with earlier studies that have identified similar challenges in working with social media data. These challenges highlight critical areas for future research, emphasizing the need for:

1) Advanced sentiment analysis tools: The development of tools capable of accurately interpreting complex emotions expressed through various content types and linguistic nuances on social media.

2) Improved data extraction techniques: The exploration of efficient methods adaptable to the dynamic nature of social media platforms and APIs, while ensuring user privacy and data compliance [16], [23].

The discussion section highlights the efficacy of the WHS method in providing accurate and personalized tourist attraction recommendations. However, it also underscores the challenges in sentiment analysis and data extraction. Future research should focus on developing advanced sentiment analysis tools and improving data extraction techniques to further enhance the performance and reliability of hybrid recommendation systems.

VI. Conclusions and Future Directions

This section provides a summary of the research conclusions and suggests future directions, highlighting key findings, limitations, implications, and areas for further investigation.

A. Conclusions

The WHF model demonstrates significant potential for personalized recommendations in tourism, outperforming CF and CB models in Precision, Recall, and F1-score. Its ability to align recommendations with users' emotions and preferences highlights the model's superiority. This success has broader implications for recommendation systems across various sectors, where aligning with user emotions and preferences can enhance satisfaction and engagement through personalized experiences.

B. Limitations

The extraction of large volumes of data from social media is time-consuming and requires careful handling, slowing the process. Additionally, sentiment analysis faces challenges when dealing with abbreviations and slang, complicating accurate interpretation.

C. Future Directions

To further enhance the model's capabilities and expand its applications, the following strategies are proposed:

1) Advanced hybrid data preprocessing techniques: Implementing sophisticated hybrid data preprocessing methods to improve model efficiency and performance across various databases. This will facilitate more accurate comparisons and refinements, leading to superior recommendation accuracy.

2) Image-based preference analysis: Utilizing image-based approaches to analyze user preferences more accurately for travel products. Integrating visual data will enable the recommendation system to better understand and predict user interests.

By adopting these strategies, the WHF model can address current challenges in sentiment analysis and data extraction, thereby advancing its personalization and accuracy. These enhancements have the potential to transform recommendation systems not only in tourism but also across other sectors. Ongoing development and the integration of advanced techniques will ensure that recommendation systems continue to evolve, providing increasingly personalized and effective solutions to meet diverse user needs.

Acknowledgment

I sincerely thank the Ministry's studentship division for their generous scholarship, which has been vital to my doctoral research.

References

[1] S. Naseem, "The Role of Tourism in Economic Growth: Empirical Evidence from Saudi Arabia," Economies, vol. 9, no. 3, Art. no. 3, Sep. 2021, doi: 10.3390/economies9030117.

[2] L. S. Budovich, "The impact of religious tourism on the economy and tourism industry," HTS Teol. Stud. Theol. Stud., vol. 79, no. 1, May 2023, doi: 10.4102/hts.v79i1.8607.

[3] Shreeraksha Shankar and Hampesh K S, "Impact of Travel Blogs and Vlogs in Social Media on Tourism," Int. J. Eng. Technol. Manag. Sci., vol. 7, no. 1, pp. 16–22, Feb. 2023, doi: 10.46647/ijetms.2023.v07i01.004.

[4] M. Alonazi, "Analyzing Sentiment in Terms of Online Feedback on Top of Users' Experiences," Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 11, 2023, doi: 10.14569/IJACSA.2023.0141114.

[5] V. Balakrishnan, V. Govindan, N. I. Arshad, L. Shuib, and E. Cachia, "Facebook user reactions and emotion: an analysis of their relationships among the online diabetes community," Malays. J. Comput. Sci., pp. 87–97, Dec. 2019, doi: 10.22452/mjcs.sp2019no3.6.

[6] J. Kim and C. Stavrositu, "Feelings on Facebook and their correlates with psychological well-being: The moderating role of culture," Comput. Hum. Behav., vol. 89, pp. 79–87, Dec. 2018, doi: 10.1016/j.chb.2018.07.024.

[7] D. D. Albesta, M. L. Jonathan, M. Jawad, O. Hardiawan, and D. Suhartono, "The impact of sentiment analysis from user on Facebook to enhanced the service quality," Int. J. Electr. Comput. Eng., vol. 11, no. 4, Art. no. 4, Aug. 2021, doi: 10.11591/ijece.v11i4.pp3424-3433.

[8] R. N. Mule and S. S. Mulik, "TRS – A rule based personalized tourism recommender system," Int. J. Res. Appl. Sci. Eng. Technol., vol. 10, no. 10, pp. 1280–1285, Oct. 2022, doi: 10.22214/ijraset.2022.47172.

[9] W. Xia, "Digital Transformation of Tourism Industry and Smart Tourism Recommendation Algorithm Based on 5G Background," Mob. Inf. Syst., vol. 2022, p. e4021706, Sep. 2022, doi: 10.1155/2022/4021706.

[10] D. Roy and M. Dutta, "A systematic review and research perspective on recommender systems," J. Big Data, vol. 9, no. 1, p. 59, May 2022, doi: 10.1186/s40537-022-00592-5.

[11] Z. Zhang et al., "Scholarly recommendation systems: a literature survey," Knowl. Inf. Syst., vol. 65, no. 11, pp. 4433–4478, Nov. 2023, doi: 10.1007/s10115-023-01901-x.

[12] G. Hui, C. Mang, Z. LiQing, and X. ShiKun, "Research on Personalized Recommendation Algorithms Based on User Profile," Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 3, 2024, doi: 10.14569/IJACSA.2024.0150330.

[13] S. Bhattacharya, D. Sarkar, D. K. Kole, and P. Jana, "Chapter 9 - Recent trends in recommendation systems and sentiment analysis," in Advanced Data Mining Tools and Methods for Social Computing, S. De, S. Dey, S. Bhattacharyya, and S. Bhatia, Eds., in Hybrid Computational Intelligence for Pattern Analysis. , Academic Press, 2022, pp. 163–175. doi: 10.1016/B978-0-32-385708-6.00016-3.

[14] M. Rhanoui, M. Mikram, S. Yousfi, A. Kasmi, and N. Zoubeidi, "A hybrid recommender system for patron driven library acquisition and weeding," J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 6, pp. 2809–2819, Jun. 2022, doi: 10.1016/j.jksuci.2020.10.017.

[15] M. Kshour, M. Ebrahimi, S. Goliaee, and R. Tawil, "New recommender system evaluation approaches based on user selections factor," Heliyon, vol. 7, no. 7, p. e07397, Jul. 2021, doi: 10.1016/j.heliyon.2021.e07397.

[16] F. Ricci, D. Massimo, and A. De Angeli, "Challenges for Recommender Systems Evaluation," in CHItaly 2021: 14th Biannual Conference of the

Italian SIGCHI Chapter, in CHItaly ’21. New York, NY, USA: Association for Computing Machinery, Jul. 2021, pp. 1–5. doi: 10.1145/3464385.3464733.

[17] M. Hu, H. Li, H. Song, X. Li, and R. Law, “Tourism demand forecasting using tourist-generated online review data,” Tour. Manag., vol. 90, p. 104490, Jun. 2022, doi: 10.1016/j.tourman.2022.104490.

[18] A. R. Alaei, S. Becken, and B. Stantic, “Sentiment Analysis in Tourism: Capitalizing on Big Data,” J. Travel Res., vol. 58, no. 2, pp. 175–191, Feb. 2019, doi: 10.1177/0047287517747753.

[19] I. Mazlan, N. Abdullah, and N. Ahmad, “Exploring the Impact of Hybrid Recommender Systems on Personalized Mental Health Recommendations,” Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 6, 2023, doi: 10.14569/IJACSA.2023.0140699.

[20] X. Zheng, Y. Luo, L. Sun, J. Zhang, and F. Chen, “A tourism destination recommender system using users’ sentiment and temporal dynamics,” J. Intell. Inf. Syst., vol. 51, no. 3, pp. 557–578, Dec. 2018, doi: 10.1007/s10844-018-0496-5.

[21] Y. A. Akbar, Z. A. Baizal, and A. T. Wibowo, “Tourism Recommender System using Weighted Parallel Hybrid Method with Singular Value Decomposition,” Indones. J. Comput. Indo-JC, vol. Vol. 6 No. 2, pp. 53-64 Pages, Sep. 2021, doi: 10.34818/INDOJC.2021.6.2.579.

[22] K. X. Li, M. Jin, and W. Shi, “Tourism as an important impetus to promoting economic growth: A critical review,” Tour. Manag. Perspect., vol. 26, pp. 135–142, Apr. 2018, doi: 10.1016/j.tmp.2017.10.002.

[23] Y. M. Arif, D. D. Putra, D. Wardani, S. M. S. Nugroho, and M. Hariadi, “Decentralized recommender system for ambient intelligence of tourism destinations serious game using known and unknown rating approach,” Heliyon, vol. 9, no. 3, p. e14267, Mar. 2023, doi: 10.1016/j.heliyon.2023.e14267.

[24] Z. Abbasi-Moud, S. Hosseinabadi, M. Kelarestaghi, and F. Eshghi, “CAFOB: Context-aware fuzzy-ontology-based tourism recommendation system,” Expert Syst. Appl., vol. 199, p. 116877, Aug. 2022, doi: 10.1016/j.eswa.2022.116877.

[25] M. Chu, Y. Chen, L. Yang, and J. Wang, “Language interpretation in travel guidance platform: Text mining and sentiment analysis of TripAdvisor reviews,” Front. Psychol., vol. 13, Oct. 2022, doi: 10.3389/fpsyg.2022.1029945.

[26] N. A. K. M. Haris, S. Mutalib, A. M. A. Malik, S. Abdul-Rahman, and S. N. K. Kamarudin, “Sentiment classification from reviews for tourism analytics,” Int. J. Adv. Intell. Inform., vol. 9, no. 1, Art. no. 1, Mar. 2023, doi: 10.26555/ijain.v9i1.1077.

[27] S. Huang, X. Wu, X. Wu, and K. Wang, “Sentiment analysis algorithm using contrastive learning and adversarial training for POI recommendation,” Soc. Netw. Anal. Min., vol. 13, no. 1, p. 75, Apr. 2023, doi: 10.1007/s13278-023-01076-x.

[28] Z. Yuan, “Big Data Recommendation Research Based on Travel Consumer Sentiment Analysis,” Front. Psychol., vol. 13, Feb. 2022, doi: 10.3389/fpsyg.2022.857292.

[29] M. Balfaqih, “A Hybrid Movies Recommendation System Based on Demographics and Facial Expression Analysis using Machine Learning,” Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 11, 2023, doi: 10.14569/IJACSA.2023.0141177.

[30] R. C K and S. K C, “Weighted hybrid model for improving predictive performance of recommendation systems using ensemble learning,” Indian J. Comput. Sci. Eng., vol. 13, no. 2, pp. 513–524, Apr. 2022, doi: 10.21817/indjcse/2022/v13i2/221302133.

[31] S. Shah, Y. Raisinghani, and N. Gandhi, “Weighted Hybrid Recommendation System Using Singular Value Decomposition and Cosine Similarity,” in Soft Computing and its Engineering Applications, K. K. Patel, G. Doctor, A. Patel, and P. Lingras, Eds., in Communications in Computer and Information Science. Cham: Springer International Publishing, 2022, pp. 367–381. doi: 10.1007/978-3-031-05767-0_29.

[32] M. Elahi, D. Khosh Kholgh, M. S. Kiarostami, M. Oussalah, and S. Saghari, “Hybrid recommendation by incorporating the sentiment of product reviews,” Inf. Sci., vol. 625, pp. 738–756, May 2023, doi: 10.1016/j.ins.2023.01.051.

[33] L. Chen, J. Cao, W. Liang, J. Wu, and Q. Ye, “Keywords-enhanced Deep Reinforcement Learning Model for Travel Recommendation,” ACM Trans. Web, vol. 17, no. 1, p. 5:1-5:21, Dec. 2022, doi: 10.1145/3570959.

[34] M. Kolahkaj, A. Harounabadi, A. Nikravanshalmani, and R. Chinipardaz, “A hybrid context-aware approach for e-tourism package recommendation based on asymmetric similarity measurement and sequential pattern mining,” Electron. Commer. Res. Appl., vol. 42, p. 100978, Jul. 2020, doi: 10.1016/j.elerap.2020.100978.

[35] L. Chen, J. Cao, Y. Wang, W. Liang, and G. Zhu, “Multi-view Graph Attention Network for Travel Recommendation,” Expert Syst. Appl., vol. 191, p. 116234, Apr. 2022, doi: 10.1016/j.eswa.2021.116234.

[36] “Graph API - Documentation,” Meta for Developers. Accessed: Nov. 12, 2023. [Online]. Available: https://developers.facebook.com/docs/graph-api/.

[37] M. Burkhardt, A. Helmond, T. Seitz, and F. van der Vlist, “The evolution of Facebook’s GRAPH API,” AoIR Sel. Pap. Internet Res., Oct. 2020, doi: 10.5210/spir.v2020i0.11185.

[38] W. Phatthiyaphaibun et al., “PyThaiNLP: Thai Natural Language Processing in Python,” in Proceedings of the 3rd Workshop for Natural Language Processing Open Source Software (NLP-OSS 2023), L. Tan, D. Milajevs, G. Chauhan, J. Gwinnup, and E. Rippeth, Eds., Singapore: Association for Computational Linguistics, Dec. 2023, pp. 25–36. doi: 10.18653/v1/2023.nlposs-1.4.

[39] Y. Qian, Y. Zhang, X. Ma, H. Yu, and L. Peng, “EARS: Emotion-aware recommender system based on hybrid information fusion,” Inf. Fusion, vol. 46, pp. 141–146, Mar. 2019, doi: 10.1016/j.inffus.2018.06.004.

[40] T. O. Hodson, “Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not,” Geosci. Model Dev., vol. 15, no. 14, pp. 5481–5487, Jul. 2022, doi: 10.5194/gmd-15-5481-2022.

[41] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” Mach. Learn. PYTHON, vol. 12, pp. 2825–2830, 2011.

[42] A. Pramarta and A. Baizal, “Hybrid recommender system using singular value decomposition and support vector machine in bali tourism,” JIPI J. Ilm. Penelit. Dan Pembelajaran Inform., vol. 7, no. 2, Art. no. 2, May 2022, doi: 10.29100/jipi.v7i2.2770.