Development and Evaluation of *Viajefy*: A Tourism Information System using TF-IDF Algorithm

Arpee C. Arruejo* and Richard C. Arruejo
College of Communication and Information Technology
University of Northern Philippines
Heritage City of Vigan, Ilocos Sur, 2700 Philippines
**arpee.callejo@unp.edu.ph*

Date received: May 23, 2024
Revision accepted: March 20, 2025

Abstract

The study aimed to design, develop, and evaluate a tourism information system called "Viajefy," incorporating the TF-IDF algorithm and assessing its model performance. It employed Feature-Driven Development as the software development method. It utilized the Cross-Industry Standardized Process for Data Mining for the data mining process of the TF-IDF algorithm, serving as its recommender agent feature. The confusion matrix evaluation tool was used to assess the algorithm's performance, yielding an accuracy of 97%, a precision of 93%, and a recall of 90%. Results showed that the recommender agent of the software application was proven reliable based on the algorithm's performance criteria in terms of accuracy, precision, and recall, and the system received a "Very Highly Acceptable" rating of 4.74. This software application is one of the first studies along tourism information systems for Ilocos Sur, Philippines, to integrate a recommender agent to help the Provincial Government of Ilocos Sur advertise attractions and establishments to be managed by the said government, where one of the Seven Wonders of the World, Vigan City, is situated.

Keywords*:* confusion matrix, CRISP-DM, machine-learning, natural language processing, text mining

1. Introduction

The COVID-19 pandemic has greatly impacted the global economy. Many people lost their jobs, numerous companies halted operations, and tragically, many lives were lost. The coronavirus (COVID-19) crisis was deemed the most severe economic shock the United States has faced (Meyer *et al*., 2022). In the Philippines, Filipinos are resilient and have consistently found ways to survive. As a result, the buying and selling of products online, as well as other opportunities, have become increasingly prevalent. The Philippines must recover from the economic losses during the long quarantine period. While these issues increased, they also opened opportunities for many. These have generated a massive amount of data used by internet users, as well as in tourism and e-commerce. Therefore, data processing is necessary to derive insights for informed decision-making. Filtering information on the web for websites and information systems is a significant feature among information systems. As stated in a previous study by Suryachandra and Reddy (2020), well-informed decisions are required in all professional fields. Complete knowledge is needed to evaluate customer reviews for all big data applications. Gurumoorthy and Suresh (2020) also claimed that nowadays, several e-commerce giants, including Amazon, Flipkart, and eBay, have been built at the forefront using machine learning (ML), which certainly manages the world.

Ilocos Sur is one of the Philippines' most famous tourist destinations, and one of the Seven Wonders of the World is located at the heart of Ilocos Sur, the Heritage City of Vigan. According to Quebral (2020), tourism in the province increased significantly in 2015 because of Vigan's inclusion as one of the New Seven Wonders Cities in late 2014. The area lacks a dedicated tourism website to promote the local tourism industry. Many tourism websites in the Philippines promote the tourism industry; however, no specific website focuses on promoting local tourism in the province. Moreover, Auliarahman *et al*. (2021) stated that having a wide variety of tourist attractions in a location might overwhelm travelers as they struggle to identify activities that align with their preferences. Thus, this study developed a system integrated with a recommender agent to promote the tourism industry in the province.

The recommender system can use the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm. The studies of Johari and Laksito (2021) and Husin *et al*. (2023) used TF-IDF to filter recommended items for users. This algorithm weighs the frequency of words on a specific item against other items, which can be added to the recommendation list of another item in a system's database.

Social networks like Facebook, Twitter, Instagram, and LinkedIn, as well as e-commerce sites like Amazon.com, Best Buy, and Spotify, and online dating sites like Tinder, utilize recommender information retrieval technology. Additionally, popular sites such as YouTube, Spotify, and Netflix employ this technology. LinkedIn, a professional social networking site, also utilizes a recommender system. These technologies are increasingly common in industries like e-commerce, music, and video streaming and are now entering the educational space (Algarni and Sheldon, 2023). They provide tailored recommendations, making them highly relevant for course recommendations. Moreover, the TF-IDF algorithm thoroughly evaluates a word's significance inside a text or group of texts (Liang and Niu, 2022).

In the Philippines, a recommender system is widely used in various applications, such as recommending clothes, books, movies, and more (Mendez and Bulanadi, 2020). In addition, Natividad *et al*. (2019) developed a fuzzy-based career recommender system for senior high school students in the K-12 education system in the Philippines, concluding that it represents significant research work in the new era of education.

In the tourism sector, Hassannia *et al*. (2019) developed an autonomous application to enhance online shopping and introduced the creation of intelligent software by implementing a powerful recommendation system. Thus, this study could enhance the tourism sector in Ilocos Sur, Philippines, to provide a more personalized experience for browsing tourism sites in the region.

A recommender system is a machine learning algorithm that utilizes the TF-IDF algorithm to weigh the frequency of words between items, thereby making recommendations. As one of the most visited tourist destinations in the Philippines, developing a recommender system for the tourism industry in Ilocos Sur is a suitable approach. The system utilizes a web application to suggest and recommend tourist attractions and other establishments to tourists and travelers within the province. This way, products and tourist attractions will be promoted, and new job opportunities will be created for Ilocanos.

This study aimed to develop a tourism information system and employ a text-mining algorithm for the province of Ilocos Sur's tourism sector. Specifically, it determined the strengths and weaknesses of the existing tourism information system, developed an online tourism information system for the Province of Ilocos Sur and identified the algorithm to be used, evaluated the level of accuracy, precision, and recall of the recommender agent of the tourism information system and determine the acceptability of the Ilocos Sur Viajefy in terms of the following UTAUT2 tool.

2. Methodology

2.1 Software Development Methodology

This study utilized the Feature Driven Development (FDD) methodology. It is built on pre-established software development standards, includes five simple operations, and is designed to be simple to produce, versatile, and manageable. The FDD's primary goal is to design and develop software per feature on time (Andry *et al*., 2020). It is one of the agile methodologies in software development that focuses on features. FDD is a suitable software method because it has customer-centric processes for developing software to fit the user requirements and specifications (Costales *et al*., 2020). Figure 1 shows the diagram of deliverables to be done to align with FDD.

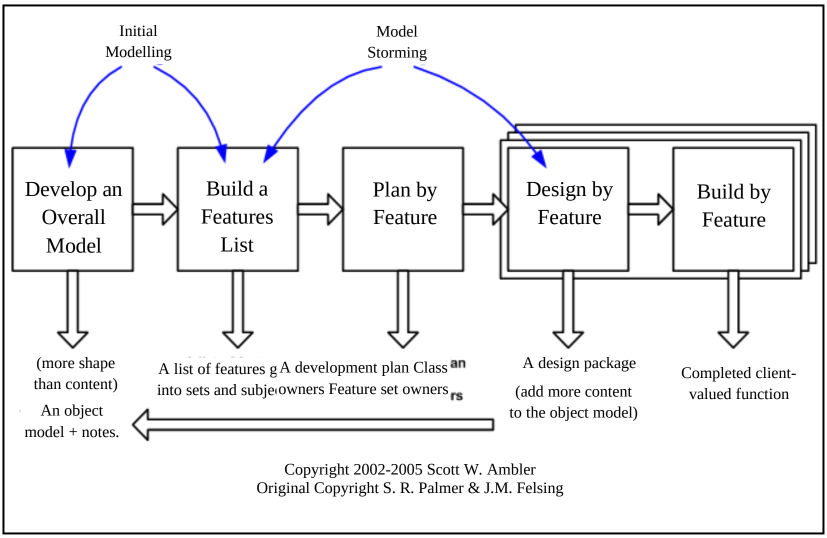

Figure 1. Feature-Driven Development

2.1.1 Develop an overall Model

This study developed the system's overall prototype and all the necessary variables. This stage is classified as the beginning of the application's development to gather the required data for the system.

2.1.2 Build Feature List

The application's features were identified and listed in this phase. A list of the various characteristics of the recommender system was planned and prepared. The stage establishes a feature list that serves as a criterion for the following step based on client requests (Andry *et al*., 2020). Lists of primary features and primary feature activities are the outcomes of this procedure. Since the development is based on customer requests, this paradigm is customer-centric.

2.1.3 Plan by Features

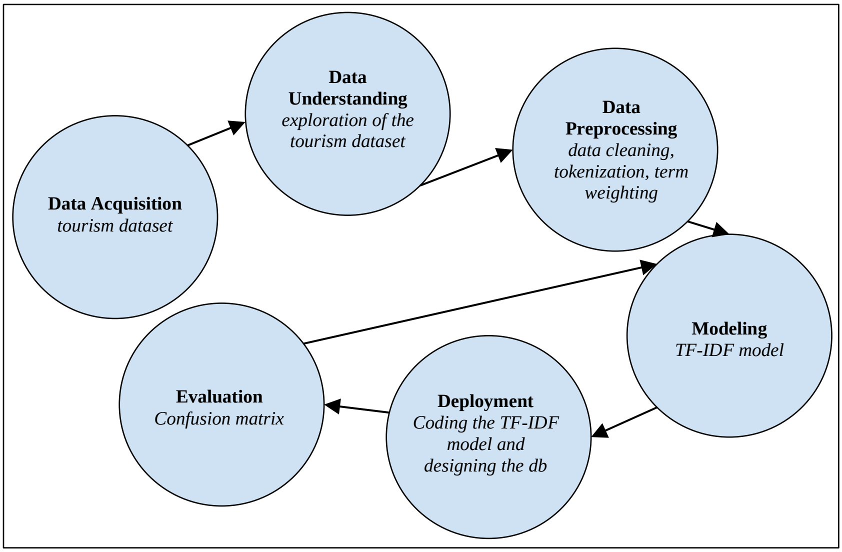

Figure 2. Process of the TF-IDF

Concrete blueprints for each identified feature were created. The strategies from the earlier phases were coded and implemented. At this stage, detailed features were received and developed based on requirements and preferences. The Cross-Industry Standardized Process for Data Mining (CRISP-DM) framework was incorporated for the TF-IDF algorithm. Plotnikova *et al*. (2022) state that organizations are implementing standardized processes, particularly CRISP-DM, to manage data mining projects effectively for sustained benefits.

The CRISP-DM framework was used to identify the required data and model. Its phases include business understanding (1), data understanding (2), data preparation (3), modeling (4), evaluation (5), and deployment and monitoring (6) (Figure 2). Similar to Lamarca (2020), the processes of business understanding, data understanding, and data preparation were applied to process, clean, and consolidate data for selecting the appropriate data mining algorithm.

2.1.3.1 Business Understanding

In this phase, the data for recommendations was identified. Training data sets were collected to build the similarity between each item in the system's database, such as names of attractions, natural features, hotels, and restaurants.

2.1.3.1.1 Data Acquisition

For the system's algorithm, the CRISP-DM process was utilized. Various online tourism information from the City Tourism Office, such as attractions, hotels, and restaurants, was collected in .csv format. This dataset was used for the TF-IDF model of the developed system. Pakpahan *et al*. (2023) demonstrated that data in .csv format can be collected through web scraping from e-commerce websites.

2.1.3.2 Data Understanding

With the collected data, further exploration was conducted to assess its quality for use in the TF-IDF model.

2.1.3.3 Data Preprocessing

During the data preparation stage, the dataset was ensured to be ready for mining. The data collection underwent standardization and CSV file transformation. Tags were assigned to classify attractions into specific categories, and similar attractions were grouped accordingly. Before incorporation into the system, the descriptions were preprocessed to clean and transform the data.

2.1.3.3.1 Data Cleaning

Stop words, punctuation, and special characters were removed to clean and preprocess the text. Hosseinzadeh *et al*. (2021) supported that data cleaning aims to remove errors, resolve inconsistencies, and convert data into a standardized format for processing.

2.1.3.3.2 Tokenization

Tokenization is vital for weighting keywords in model items. It splits item descriptions from the datasets, resulting in single-word tokens of importance based on TensorFlow (Prastyo *et al*., 2020; Pakpahan *et al*., 2023). In the system's development, single words were assigned to each tourist attraction, hotel, and restaurant for the recommender agent. The recommendation system was created by structuring the database schema and the system's architectural design. According to IBM (2020), algorithms may be integrated into a system either through a database or programming codes.

2.1.3.3.3 Design by Feature

It includes the processes and tools used in developing the system. The front-end and back-end frameworks of the system include PHP, HTML, and MySQL. All user interfaces were plan to provide an enhanced user experience on the system.

2.1.3.3.4 Build by Feature

This is the final stage, where the application's plan and design features are created. Using the previous step as a guide, the details of the features were transformed into a functional system for customers with the help of various tools and processes. The CRISP-DM processes—modeling, evaluation, and deployment—were also incorporated into the system.

2.1.3.3.5 Modelling

The TF-IDF model was used because it is the most precise and concise model for such a system based on literature. Churches, heritage, museums, nature, historical, hotels, and restaurants were identified and got the most labels from the data sets. The agent filters items from the database to be recommended for users. In this way, the system helps users decide what to choose from.

The Term Frequency-Inverse Document Frequency (TF-IDF) algorithm was chosen for text-based classification in natural language processing. Liang and Niu. (2022) highlighted its effectiveness in assessing a word's significance within a text or class of texts. The formula for TF-IDF is explained below. Using Equations 1, 2, and 3, item classification for the system's recommender agent was determined.

$$TFij = \frac{Nij}{Dj} \quad (1)$$

where TFij denotes the TF value of the i-th word for the j-th document,

$$IDFi = lg\ (\frac{D}{Ni}) \quad (2)$$

where IDFi denotes the IDF value of the i-th word, a

$$TF - IDFij = TFij.IDFi \quad (3)$$

and TF-IDFij denotes the TF-IDF value of the i-th word for the j-th document. D is the total number of all documents, Dj is the j-th document, nij is the number of occurrences of the i-th word in the j-th document, and Ni is the number of occurrences of the i-th word in all documents.

2.1.3.3.6 Item Recommendation

The "You might like this" part of the developed system represents the recommendations. Recommendations are presented and are ranked by the distance of the similarity of each item in the documents.

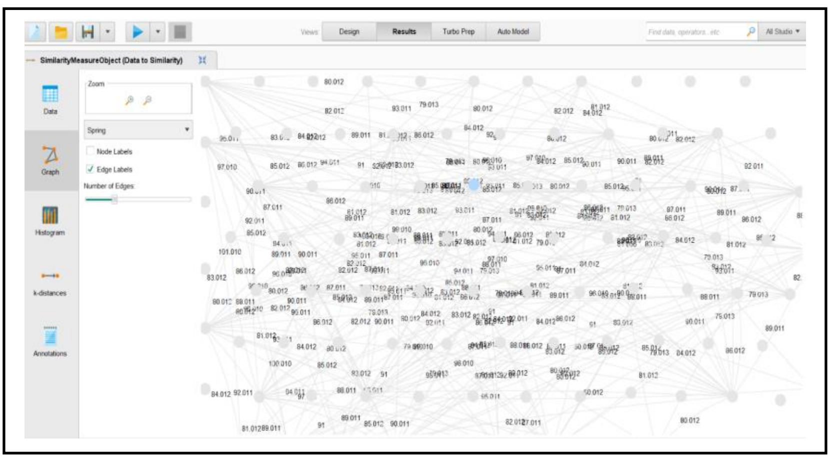

Figure 3. Distance of each item

Figure 3 shows the results on the distance of every item. These results of the Rapid Miner show the similarity of each item to the other of the TF-IDF algorithm. The recommendation agent is created for the system by structuring the database schema and instance of the system architectural design. The classification was done through tags and keywords inserted in the system. According to Watanobe *et al*. (2023), algorithms may be integrated in the form of a database or through the programming codes of the system.

2.1.3.3.7 Deployment

In this phase, tools were identified to integrate the model into the system. MySQL was used as the backend to construct the system's database, providing a storage solution for user data. Essential components for the system model included PHP, CSS, and TF-IDF scripts for algorithm deployment. Kurniawan *et al*. (2020) highlighted the use of two models: RapidMiner Server or application development through a selected programming language. Previous studies have shown that VB.net and PHP are commonly used for deployment. Thus, PHP and MySQL were chosen for the database.

2.1.3.3.8 Evaluation

In this stage, the model's accuracy, precision, and recall for the recommender system agent were evaluated using a confusion matrix. Singh *et al*. (2021) stated that a confusion matrix provides insights into a classification algorithm's performance. Accuracy, precision, and recall were used to validate the model's output, aligning with Suryati *et al*. (2023), who emphasized the importance of model evaluation for performance assessment. Additionally, precision and recall are widely recommended as key metrics for classifier performance. These measures are computed using Equation 4, which validates the accuracy of the recommender agent's recommendations.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

The accuracy of the recommender agent is validated through the confusion matrix. Equation 4 is the computation of the results of the model's accuracy where True-Positive (TP) added to True Negative (TN) are divided into all data in the database of a specific category. The model should satisfy 80% to prove the model of the recommender agent. Also, Itahashi *et al*. (2024) cited that acceptable accuracy percentages vary by domain. In consumer applications, 70% – 85% accuracy is acceptable (Table 1); in high-risk fields like healthcare or finance, accuracy should exceed 90%.

Table 1. Accuracy of Model Performance

Range of Percentage	Description
>= .85 & <= .1	High Level of Performance
> .70 & < .85	Acceptable Performance
< .70	Not Acceptable Performance

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

Equation 5 was used to measure the precision of the recommender agent of the system. This is computed through the division of the True-Positive (TP) items into True-Positive (TP) and False-Negative (FN) items of the recommender agent. A 78% recall must be satisfied to prove the model of the system.

$$Recall = \frac{TP}{TP+FN} \tag{6}$$

For recall (Equation 6), the recommender agent is calculated with the correctly predicted items (TP) to the correctly predicted items (TP) and incorrectly predicted items (FN) on the database of the system developed. A model should be above 75% as the standard acceptable prediction recall. For precision, an acceptable percentage is typically above 75% (Table 2) for most models (Saito and Rehmsmeier, 2018). Recall should also exceed 75%, especially when missing true positives are costly.

Table 2. Precision and Recall of Model Performance

Range of Percentage	Description
>= .75	Acceptable Performance
< .75	Not Acceptable Performance

With this process, there is an avenue for users to test the system through the Unified Theory of Acceptance and Use of Technology (UATAUT2). According to Dwivedi *et al*. (2019), the acceptability tool is an alternative theoretical model for explaining the acceptance and use of information system (IS) and information technology (IT) innovations. Indicators of this tool include performance expectancy, effort expectancy, social influence, hedonic motivation and behavioral intention.

The first indicator, performance expectancy, is the degree to which an individual believes that using the system improves his/her actions. According to Arain *et al*. (2019), perceived usefulness and outcome expectations are key factors in determining technology acceptance and performance expectancy in higher education.

Another indicator is effort expectancy, which is the perception that using the system is effortless. Almaiah and Mulhem (2019) supported the idea that users' perceptions of ease of use are closely tied to their willingness to continue using the system, with social influence playing a role, especially among students. Effort expectancy refers to how simple it is to use an information system.

Social influence is another indicator in AUTUT2. Social Influence is the extent to which an individual perceives that it is important for others to believe that s/he should adopt technology. Zacharis and Nikolopoulou (2022) found that social influence significantly affects individuals' behavioral intentions regarding mobile banking adoption. This effect is stronger in proactive individuals who perceive their social network as supportive of new technologies.

Another indicator is hedonic motivation. Tamilmani *et al.* (2019) defined hedonic motivation as "*the fun or pleasure derived from using technology, and it is an important determinant of consumer's technology acceptance and use"*. Hedonic motivation within UTAUT2 influences users' engagement with technology systems. Additionally, Vimalkumar *et al*. (2021) also explored hedonic motivation as a key factor that influences various technologies.

Lastly is the behavioral intention of AUTUT2. According to Fitrianie *et al*. (2021), the UTAUT2 model states that people's behavioral intention is one important mediating factor for behavior under free will. Thus, this is also an important factor to consider in the acceptance of the system. Table 3 shows the evaluation's range, descriptive equivalent, and descriptive interpretation.

Table 3. Unified Theory of Acceptance and Use of Technology

Range	Descriptive Equivalent	Descriptive Interpretation
4.51- 5.00	Very Strongly Agree	Very Highly Acceptable
3.51 - 4.50	Strongly Agree	Highly Acceptable
2.51 - 3.50	Agree	Acceptable
1.51 - 2.50	Disagree	Fairly Acceptable
1.00 - 1.50	Strongly Disagree	Not Acceptable

3. Results and Discussion

3.1 On Strengths and Weaknesses of the Existing Technologies for Ilocos Sur, Philippines along tourism

A tourism information system for the province was needed. Currently, the only existing technologies for the tourism industry in Ilocos Sur are the mobile application *Tour Ilocos Sur* and a tourism webpage on the official website of the Provincial Government of Ilocos Sur. The strengths and weaknesses of these solutions were assessed through interviews and internet research. Additionally, limitations in their features and functions were identified.

The strengths of the existing tourism information systems were identified, focusing on the *Tour Ilocos Sur* mobile application. Key strengths include user accessibility, device integration, and speed. Users can quickly access the app on smartphones or tablets and interact with its features. Additionally, mobile apps can integrate device functions such as the camera, microphone, contacts, and GPS. However, despite these advantages, the mobile application operates independently of the Provincial Tourism Office, as noted by the current Tourism Officer. Consequently, no dedicated tourism information system exists for the province.

Likewise, weaknesses were also identified. These are tourism information system unavailability, search engine invisibility, cross-platform incompatibility, limited sharing and linking, and search engine optimization (SEO) unavailability. The unavailability of a tourism information system in the province makes the tourism industry in Ilocos Sur invisible to an online audience. Search engine invisibility allows users to find and access attractions with difficulty through search queries. Effective SEO tactics can improve the website's visibility and organic traffic. A cross-platform compatible with any device means promoting tourism is more accessible and visible to prospective travelers worldwide. Another limitation of apps is the ability to share and link tourist attractions.

All platforms, including desktop computers, laptops, tablets, and smartphones, can browse websites. Websites that use responsive web design can adjust to various screen sizes and deliver consistent user experience. That is why this type of design was considered. Lastly is SEO. Having a tourism information system might benefit from specialized SEO tactics and optimization for search engines to promote the tourism industry in Ilocos Sur.

While tourism promotion in the province remains a top priority, an information system for promoting attractions requires a recommendation feature similar to those used in platforms like Lazada, Shopee, Amazon, Netflix, Spotify, and YouTube. Interviews with research participants highlighted limitations in existing systems, particularly the absence of a recommendation agent to suggest attractions, establishments, and products to users of both the mobile application and the tourism webpage. Yalçın (2021) emphasized that recommender systems aim to provide relevant suggestions to users, fostering trust and supporting long-term business objectives. Similarly, Alconis and Aquino (2022) suggested that a tourism website could be beneficial for Metro Vigan. Thus, incorporating a recommender agent could enhance user experience by offering personalized travel selections for visitors to Ilocos Sur.

3.2 Development of an online tourism information system for the Province of Ilocos Sur, Philippines incorporating the algorithm for its recommender agent.

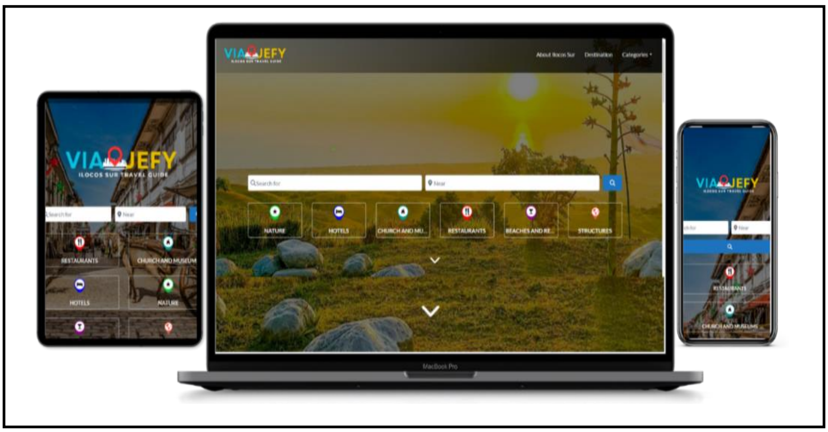

Figure 4. Ilocos Sur Philippines Viajefy: A Tourism Information System

Figure 4 shows the Tourist Information System with the recommender agent. The system assists users with the province's different tourist attractions and establishments. The system is also dynamic and responsive, working both on mobile and desktop to address the findings of the problems of the application. Features of the system include: travel guide (1), a recommender agent (2), user reviews (3), a search feature by popularity, recently added and featured (4), and geo-reference (5).

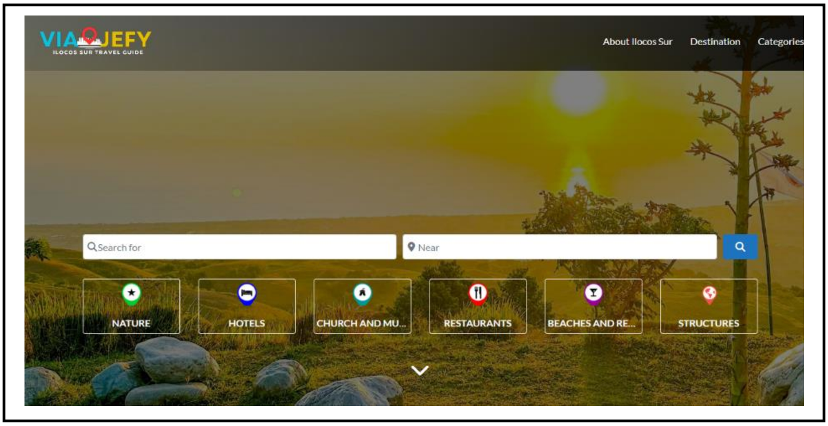

Figure 5. The Home Page of *Viajefy* Recommender Agent

Travel Guide (Figure 5)**.** Primarily, the system offers assistance to tourists. It offers a list of tourist attractions, hotels, restaurants, and others.

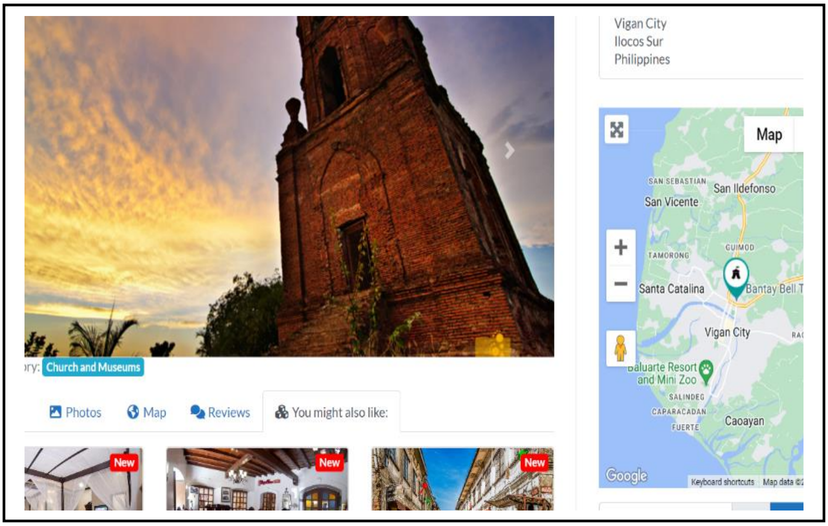

Figure 6. Recommender Agent

Recommender Agent (Figure 6). This offers a similarity feature through an algorithm that makes the system feature, recommender agent, work. It offers similar items for users to view for reviews and comments. It is embedded in the system for users to select the best options when visiting the province.

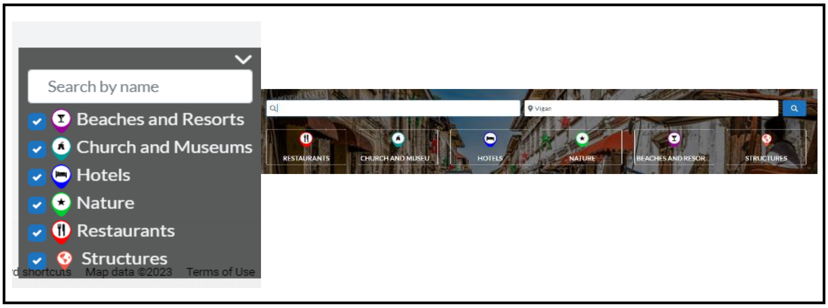

Figure 7. Search by Categories

Search (Figure 7)**.** Another feature is the search filtering by categories. Three categories are included in the search feature, which can help users search their options.

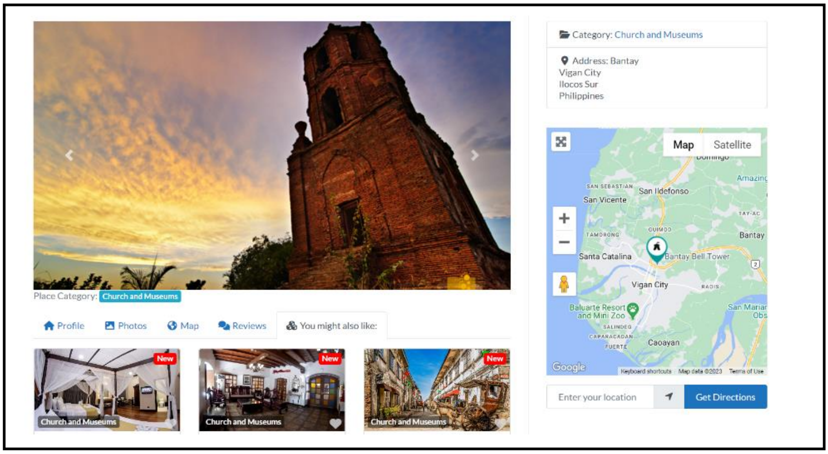

Figure 8. Geo-references

Geo Reference (Figure 8)**.** This feature primarily helps you locate the attractions and establishments in the province. This helps users find the location of the attractions and establishments in the province. This is located below the attractions and establishments. The performance of the algorithm was validated through the confusion matrix.

3.3 Evaluation on the level of accuracy, precision and recall of the recommender agent of the tourism information system.

One of the instruments is to test the model's accuracy, precision, and recall through a confusion matrix. According to the study of Reddi and Eswar (2021), confusion matrices were preferred when evaluating the effectiveness of classification algorithms. Thus, determining the accuracy, precision, and recall of the recommendation agent applied to the system.

The recommender agent achieved a 97% model performance in terms of accuracy and precision, with a 90% recall rate (Figure 9). According to the descriptive ratings, the model utilized in the study demonstrated a high level of performance and acceptable performance in the aforementioned criteria. It means that the recommender is effective and efficient and is recommended for deployment on any applications that require recommender agents, such as Viajefy.

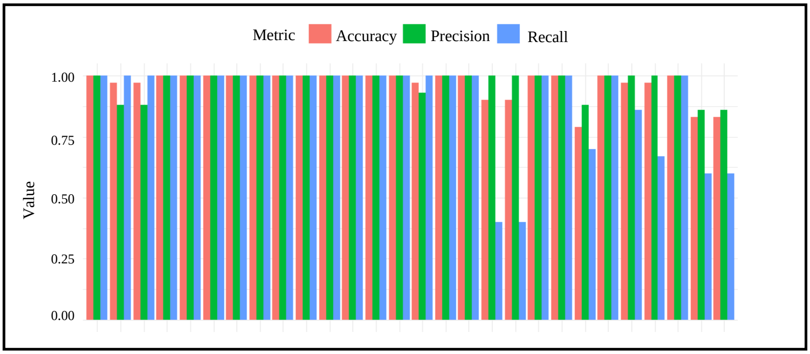

Figure 9. Model Performance on Accuracy, Precision, and Recall

3.4 Acceptability of the Ilocos Sur Viajefy in terms of Performance Expectancy, Effort Expectancy, Social Influence, Hedonic Motivation and Behavioral intention.

The level of acceptability of the developed system was tested in terms of performance expectancy, social influence, hedonic motivation, and behavioral intention (Table 4). The acceptability tool used was created by Tran and Nguyen (2021). The results are presented in Table 4. The Provincial Tourism Officer, IT expert, restaurateur, tour guide, and hotel owner evaluated the system.

Table 4. Results of the UATUT2 Evaluation on Acceptability

Indicators	Mean	Descriptive Equivalent	Descriptive Interpretation
Performance Expectancy	4.80	Very Strongly Agree	Very Highly Acceptable
Effort Expectancy	4.73	Very Strongly Agree	Very Highly Acceptable
Social Influence	4.60	Very Strongly Agree	Very Highly Acceptable
Hedonic Motivation	4.86	Very Strongly Agree	Very Highly Acceptable
Behavioral Intention	4.70	Very Strongly Agree	Very Highly Acceptable
Overall Mean	4.74	Very Strongly Agree	Very Highly Acceptable

The performance expectancy of the system was evaluated with a mean rating of 4.80, a descriptive equivalent of "Very Strongly Agree," and an interpretation of "Very Highly Acceptable." Another indicator is effort expectancy. The system was also evaluated with a mean rating of 4.73, a descriptive equivalent of "Very Strongly Agree," and an interpretation of "Very Highly Acceptable" in effort expectancy. Social influence is another indicator in UTAUT2. The system obtained a mean rating of 4.60, with a descriptive equivalent of "Very Strongly Agree" and an interpretation of "Very Highly Acceptable" in social influence. The system's hedonic motivation was evaluated with a mean rating of 4.86, a descriptive equivalent of "Very Strongly Agree" and an interpretation of "Very Highly Acceptable." Lastly is the behavioral intention of UTAUT2. The system's behavioral intention was evaluated with a mean rating of 4.70, a descriptive equivalent of "Very Strongly Agree," and an interpretation of "Very Highly Acceptable."

The system's overall mean of 4.74 has a descriptive interpretation of Very Highly Acceptable. Thus, the system is proven to be suitable for use.

4. Conclusion and Recommendation

This developed a tourist information system embedded with the TF-IDF algorithm to recommend unseen items on attractions and establishments that might need to be more familiar to tourists to address the weaknesses in the promotion of the tourism industry in the province. Likewise, the recommender agent of the software application's algorithm was proven to gain an accuracy of 97%, precision of 93%, and recall of 90%. These results show the readiness of the algorithm with the deployment of the system in the province. The system was also evaluated with a Very Highly Acceptable descriptive interpretation with an overall mean of 4.74. The system is considered highly acceptable and proven to be ready for implementation.

The research findings revealed that the developed application might help the provincial government to create more jobs and services along with tourism through the help of the recommender feature of the system. The recommender agent's accuracy, precision, and recall have been proven for use.

5. Acknowledgement

The authors are grateful to the University of Northern Philippines administration headed by Dr. Erwin F. Cadorna, to the Vice President for Research and Extension, Dr. Fatima F. Rocamora, the University Research and Development Office led by Dr. Edelyn A. Cadorna, to the University of Northern Philippines community, family, relatives, and friends who have supported the conduct of this study. This study has been realized with these people who push us to work hard. All of these are not possible without our Almighty God, Amen.

6. References

Alconis, A.L., & Aquino, L.B. (2022). The Ylocandia tourism industry in the Philippines. Science International-Lahore, 34(6), 35-39. http://www.sci-int.com/pdf/638067105997833058.pdf

Algarni, S., & Sheldon, F. (2023). Systematic review of recommendation systems for course selection. Machine Learning and Knowledge Extraction, 5(2), 560-596. https://doi.org/10.3390/make5020033

Almaiah, M.A., & Al Mulhem, A. (2019). Analysis of the essential factors affecting of intention to use of mobile learning applications: A comparison between universities adopters and non-adopters. Education and Information Technologies, 24, 1433–1468. https://doi.org/10.1007/s10639-018-9840-1

Andry, J.F. Tannady, H., & Gunawan, F.E. (2020). Purchase order information system using feature driven development methodology. International Journal of Advanced

Trends in Computer Science and Engineering, 9(2), 1107-1112. http://dx.doi.org/10.30534/ijatcse/2020/32922020

Arain, A.A., Hussain, Z., Rizvi, W.H., & Vighio, M.S. (2019). Extending UTAUT2 toward the acceptance of mobile learning in the context of higher education. Universal Access in the Information Society, 18, 659–673. https://doi.org/10.1007/s10209-019-00685-8

Auliarahman, R.M., Baizal, Z.K.A., & Ikhsan, N. (2021). Conversational recommender systems based on criticism for tourist attractions using TF-IDF. Jurnal Media Informatika Budidarma, 5(4), 1242-1248. https://doi.org/10.30865/mib.v5i4.3245

Costales, H., Callejo-Arruejo, A., & Rafanan, N. (2020). Development of a prototype application for rice disease detection using convolutional neural networks. International Journal of Emerging Trends in Engineering Research, 8 (10), 7076-7081. https://doi.org/10.48550/arXiv.2301.05528

Dwivedi, Y.K., Rana, N.P., Jeyaraj, A., Clement, M., & Williams, M.D. (2019). Re-examining the unified theory of acceptance and use of technology (UTAUT): towards a revised theoretical model. Information Systems Frontiers, 21, 719–734. https://doi.org/10.1007/s10796-017-9774-y

Fitrianie, S., Horsch, C., Beun, R.J., Griffioen-Both, F., & Brinkman, W.-P. (2021). Factors affecting user's behavioral intention and use of a mobile- phone-delivered cognitive behavioral therapy for insomnia: a small-scale UTAUT analysis. Journal of Medical Systems, 45(110). https://doi.org/10.1007/s10916-021-01785-w

Gurumoorthy, K., & Suresh, P. (2020). Supervised machine learning algorithm using sentiment analysis based on customer feedback for smart phone product. International Journal of Emerging Trends in Engineering Research, 8(8), 4456-4462. https://doi.org/10.30534/ijeter/2020/67882020

Hassannia R, Barenji, A.V., Li, Z., & Alipour H. (2019). Web-based recommendation system for smart tourism: multiagent technology. Sustainability, 11(2), 323. https://doi.org/10.3390/su11020323

Hosseinzadeh, M., Azhir, E., Ahmed, O.H., Ghafour, M.Y., Ahmed, S.H., Rahmani, A.M., & Vo, B. (2021). Data cleansing mechanisms and approaches for big data analytics: a systematic study. Journal of Ambient Intelligence and Humanized Computing, 14, 99-111. https://doi.org/10.1007/s12652-021-03590-2

Husin, M.R.M., Razak, T.R., Malik, A.M.A., Nordin, S., & Abdul-Rahman, S. (2023). Hybrid collaborative movie recommendation system. In 2023 4th International

Conference on Artificial Intelligence and Data Sciences (AiDAS), IPOH, Malaysia, 274-280. http://dx.doi.org/10.1088/1757-899X/482/1/012025

Itahashi, T., Yamashita, A., Takahara, Y., Yahata, N., Aoki, Y., Fujino, J., Yoshihara, Y., Nakamura, M., …. & Hashimoto, R. (2024). Generalizable and transportable resting-state neural signatures characterized by functional networks, neurotransmitters, and clinical symptoms in autism. Molecular Psychiatry, 30, 1466–1478. https://doi.org/10.1038/s41380-024-02759-3

Johari, M., & Laksito, A. (2021). The hybrid recommender system of the Indonesian online market products using IMDb weight rating and TF-IDF. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 5(5), 977-983. http://dx.doi.org/10.29207/resti.v5i5.3486

Kurniawan, S., Gata, W., Puspitawati, D.A., Parthama, I.K.S., Setiawan, H., & Hartini, S. (2020). Text mining preprocessing using GATA framework and RapidMiner for Indonesian sentiment analysis. In IOP Conference Series: Materials Science and Engineering, The 7th International Conference on DV-Xα Method, Semarang, Indonesia, 835, 012057. https://doi.org/10.1088/1757-899x/835/1/012057

Lamarca, B.I. (2020). Cybersecurity risk assessment of the University of Northern Philippines using PRISM approach. In IOP Conference Series: Materials Science and Engineering, The 6th International Conference on Software Engineering & Computer Systems, Pahang, Malaysia, 769, 012066. https://doi.org/10.1088/1757-899x/769/1/012066

Liang, M., & Niu, T. (2022). Research on text classification techniques based on improved TF-IDF algorithm and LSTM inputs. Procedia Computer Science, 208, 460-470. https://doi.org/10.1016/j.procs.2022.10.064

Mendez, J.S., & Bulanadi, J.D. (2020). Job matcher: A web application job placement using collaborative filtering recommender system. International Journal of Research Studies in Education, 9(2), 103-120. https://doi.org/10.5861/ijrse.2020.5810

Meyer, B.H., Prescott, B., & Sheng, X.S. (2022). The impact of the COVID-19 pandemic on business expectations. International Journal of Forecasting, 38(2), 529-544. https://doi.org/10.1016/j.ijforecast.2021.02.009

Natividad, M.C., Gerardo, B.D., & Medina, R.P. (2019). A fuzzy-based career recommender system for senior high school students in K to 12 education. In IOP Conference Series: Materials Science and Engineering, International Conference on Information Technology and Digital Applications (ICITDA 2018), Manila City, Philippines, 482, 012025. http://dx.doi.org/10.1088/1757-899X/482/1/012025

Pakpahan, D., Siallagan, V., & Siregar, S. (2023). Classification of e-commerce product descriptions with the TF-IDF and SVM methods. Sinkron: Jurnal dan Penelitian Teknik Informatika, 7(4), 2130–2137. https://doi.org/10.33395/sinkron.v8i4.12779

Plotnikova, V., Dumas, M., & Milani, F.P. (2022). Applying the CRISP-DM data mining process in the financial services industry: Elicitation of adaptation requirements. Data & Knowledge Engineering, 139, 102013. https://doi.org/10.1016/j.datak.2022.102013

Prastyo, P.H., Ardiyanto, I., & Hidayat, R. (2020). Indonesian sentiment analysis: an experimental study of four kernel functions on SVM algorithm with TF-IDF. In 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), Sakheer, Bahrain, 1-6. http://dx.doi.org/10.1109/ICDABI51230.2020.9325685

Quebral, P.N.R. (2020). Learning orientation, service innovation and performance of hotels in Ilocos Sur. Asian Journal of Education and Human Development, 1(1). https://doi.org/10.69566/ajehd.v1i1.10

Reddi, S., & Eswar, G.V. (2021). Fake news in social media recognition using modified long short-term memory network. In Intelligent Data-Centric Systems, Security in IoT Social Networks, Academic Press, 205-227. https://doi.org/10.1016/B978-0-12-821599-9.00009-1

Saito, T., & Rehmsmeier, M. (2018). The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. PLoS ONE, 10(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

Singh, P., Singh, N., Singh, K.K., & Singh, A. (2021). Diagnosing of disease using machine learning. Machine Learning and the Internet of Medical Things in Healthcare, Academic Press, India, pp. 89-111. https://doi.org/10.1016/B978-0-12-821229-5.00003-3

Suryachandra, P., & Reddy, P.V.S. (2020). A novel hybrid machine learning approach to classify the sentiment value of natural language processing in big data. International Journal of Emerging Trends in Engineering Research, 8(8), 4160-4165. https://doi.org/10.30534/ijeter/2020/21882020

Suryati, E., Styawati, S., & Aldino, A.A. (2023). Online transportation sentiment analysis using word2vec text embedding model feature extraction and support vector machine (SVM) algorithm. Journal of Technology and Information Systems. https://doi.org/10.33365/jtsi.v4i1.2445

Tamilmani, K., Rana, N.P., Prakasam, N., & Dwivedi, Y.K. (2019). The battle of brain vs. heart: a literature review and meta-analysis of “hedonic motivation” use in UTAUT2. International Journal of Information Management, 46, 222-235. https://doi.org/10.1016/j.ijinfomgt.2019.01.008

Tran, K., & Nguyen, T. (2021). Preliminary research on the social attitudes toward AI’s involvement in Christian education in Vietnam: promoting AI technology for religious education. Religions, 12(3), 208. https://doi.org/10.33365/jtsi.v4i1.2445

Vimalkumar, M., Sharma, S.K., Singh, J.B., & Dwivedi, Y.K. (2021). ‘Okay google, what about my privacy?’: user's privacy perceptions and acceptance of voice based digital assistants. Computers in Human Behavior, 120, 106763. https://doi.org/10.1016/j.chb.2021.106763

Watanobe, Y., Rahman, M.M., Amin, M.F.I., & Kabir, R. (2023) Identifying algorithm in program code based on structural features using CNN classification model. Applied Intelligence, 53, 12210-12236. https://doi.org/10.1007/s10489-022-04078-y.

Yalçın, O.G. (2021). Applied neural networks with TensorFlow 2. API oriented deep learning with Python. Apress Berkeley, CA. https://doi.org/10.1007/978-1-4842-6513-0

Zacharis, G., & Nikolopoulou, K. (2022). Factors predicting university students’ behavioral intention to use e-learning platforms in the post-pandemic normal: an UTAUT2 approach with ‘learning value’. Education and Information Technologies, 27, 12065–12082. https://doi.org/10.1007/s10639-022-11116-2