AI and Tourism in Cameroon: A Multimodal Approach for a Hybrid Recommendation System

Chelsie Dongmo, Data Science Student[1]; Melvin Fokam, Data Science Student[2]; Dave Mbiazi, PhD Student[3]

[1] Ecole Nationale Supérieure Polytechnique de Yaoundé,

[2] Département d'Informatique Université de Yaoundé1,

[3]Ecole Polytechnique de Montréal

Abstract: This article presents a multimodal approach to hybrid recommendation systems by combining data from different forms in the field of tourism in Cameroon. With the explosion of the tourism industry and the spread of social media on online platforms, users find themselves overwhelmed by a multitude of information in various formats (photos, texts, videos, etc.). The task of the latter is complicated in such a situation because it is necessary to make decisions and obtain tailor-made recommendations regarding travel-related services. In the field of e-commerce, recommendation systems (RS) are widely used to improve conversion rates by adjusting product offerings based on preferences and customer interests can also play a fundamental role in the field of tourism. If traditional RS relies exclusively on numerical evaluations to formulate recommendations, these evaluations cannot always be sufficient to offer tailor-made and precise suggestions.

Our approach combines both collaborative filtering (Col-F or User to user) filtering based on content (Con-B or Item to item) as well as the geographical area. The system we offer combines the advantages of other systems by using tools like CNN (VGG16 and ResnetT50) to analyze images of tourist sites and classify them according to tourist destinations, other deals with the analysis of user reviews with textual data using sentiment analysis techniques such as LSTM, to provide users with more accurate recommendations and finally geography is used for new users to solve the cold start problem. By integrating various recommendation methods, our system hybridizes the accuracy and relevance of its recommendations, providing a more personalized and efficient tourism experience.

Keywords : Multimodal approach, hybrid recommendation system, collaborative filtering, content-based filtering, tourism

1 Introduction

Information obesity[1] remains a challenge for Internet users in the digital and artificial intelligence era who must search and make decisions among the many options available for services such as sales, hotels but also tourist destinations. Cameroon currently has more than 200 tourist destinations, or 1/3 of the sites in Central Africa, having welcomed 1.02 million tourists in 2019. In 2021, it generated approximately 429.19 million euros in the tourism sector alone. This corresponds to 1.0% of gross domestic product and approximately 64% of all international tourism revenue in Central Africa. The most active existing sites for tourism promotion are generally pages on networks, ignoring important factors like user preference, tourism context and real-time information because it only uses one type of filtering (ConB or ColF). Then it will be ideal to minimize the time users spend searching for information and suggesting items they might not have otherwise considered, thus improving the quality of services offered by the tourist office platforms in Cameroon.

Proposing a multimodal approach for a recommender system is expected to address issues such as scalability, cold start, and fairness, which are among the major challenges. To overcome these mentioned limitations, recommendation systems have been developed to filter information and provide personalized recommendations to users based on their specific tastes and preferences [2]. The one we propose is a hybrid system that takes into account several factors namely: tourist destinations, sentiment analysis as well as geographic location for a new user before providing a recommendation.

The document is structured as follows: section 2 examines the tools and the process put in place to obtain our data; section 3 follows with the presentation of the recommendation system as well as its architecture. Section 4 shows the results obtained as well as the discussions. As for section 5, it contains the conclusion and our future work.

2 Tools and data collection

1. Tools

Python has been our programming language with libraries like Keras and Pytorch. Added to this, Resnet-50 and VGG-16 were used to extract the characteristics of the images and an RNN in particular LSTM (Long Short Term Memory) for the analysis of user opinions and the Google API for geographical areas relating to each group of images.[4]To ensure the segmentation of our different codes in order to better use them, we worked with VS Code and Jupyter Notebook and Google Collaboratory for rapid work then host the content on GitHub for those who would like to explore.

Our work was carried out on MacBook Pro with 11-core CPU, a 14-core GPU and RAM 18 GB of unified memory and 512 GB of SSD storage, then another machine Model DELL LATTITUDE E5570 with an Intel(R) Core processor (TM) i7-6820HQ CPU with a speed of 2.70GHz, 16GB RAM and 64-bit operating system.

2. Data collection

The selection of datasets was determined by several factors, including data availability and accessibility. In this work, we used data in several forms. One including images of tourist destinations which have been collected manually. These images are either bridges, falls, museums, seaside resorts, rivers, mountains etc. …The second set corresponds to the opinions of users, in particular the comments left by them in relation to the destinations visited and is presented below in the methodology.

Table.1:Number of images per website

Web sites	Number of images
CAMEROUN à Travers Nous	201
CAMER TOUR.Org	175
TOURISME DU GRAND CAMEROUN	398
TOURISME CAMEROUN	225
VISITER LE CAMEROUN AVEC MOI	226
NOUVELLE VISION CAMEOUN TOURISME	124
CAMEROUN TOURISME	200
Total	1549

Fig.1. Sample Images before pre-processing

Table.2. Sample Comments before pre-processing

	date	facebookUrl	likesCount	postTitle	profileName	text
0	2024-04-03T07:39:44.000Z	https://web.facebook.com/photo?fbid=7253332097...	0	La Rivière Lokoundje, Sud Cameroun \n\nLa riv...	MT GENIE Construction	Belles prises !
1	2024-03-15T15:44:13.000Z	https://web.facebook.com/photo?fbid=7253332097...	1	La Rivière Lokoundje, Sud Cameroun \n\nLa riv...	Edwige Frédi	J'aimerais bien y participer la prochaine fois
2	2024-03-17T05:55:31.000Z	https://web.facebook.com/photo?fbid=7253332097...	0	La Rivière Lokoundje, Sud Cameroun \n\nLa riv...	Rose Ebolo	Prévenez moi pour la prochaine fois
3	2024-03-13T15:57:00.000Z	https://web.facebook.com/photo?fbid=7253332097...	3	La Rivière Lokoundje, Sud Cameroun \n\nLa riv...	Aurelie Elouga	C'est situé où svp ?
4	2024-03-13T16:22:26.000Z	https://web.facebook.com/photo?fbid=7253332097...	1	La Rivière Lokoundje, Sud Cameroun \n\nLa riv...	Cheyrol Ada Moneyang	Ivan Mbarga donc il y'a ça dans ton village et...

3 Methodology

Our recommendation system incorporates collaborative and content-based filtering to improve the accuracy of user recommendations. The goal is to increase the reliability of the recommendation process by integrating sentiment analysis of user reviews into traditional recommendation techniques. To make a recommendation, the system goes through different phases: preprocessing to extract the characteristics of the images of tourist destinations as well as the geographical position, cleaning and contextual analysis of user opinions using LSTM, then implementation of the hybrid system by combining the result of the two previous sections.

3.1 Image analysis and processing

The preliminary work that was done was to put the images in the same size in order to better analyze them, and using pre-trained CNNs like VGG16 to extract characteristics such as entire shapes (landscapes, buildings), scene (beach, mountain, city), simple shapes, recurring patterns, etc. These extracted characteristics will enrich the representation of destinations by retaining unique visual aspects, allow finer personalization by identifying users' visual preferences, facilitate the discovery of new destinations and improve the quality and diversity of recommendations.

Then, we created a matrix where the rows represent the users and the columns the images, with values indicating the interactions between them.

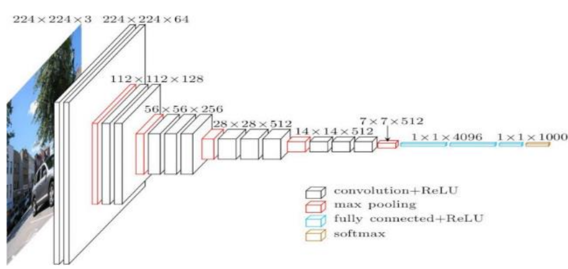

Fig2: Architecture VGG16

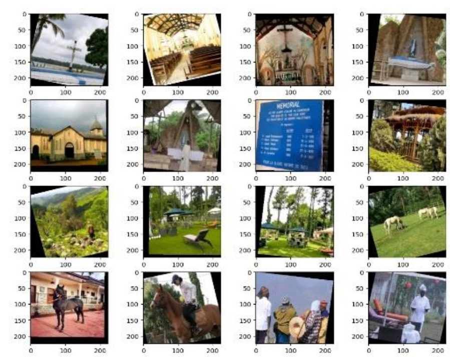

Fig3: Samples Images after pre-processing

3.2 Analysis and processing of user reviews

3.1.1 Data preprocessing

The text data was retrieved from websites and contains user inputs in total 7600 lines which will be cleaned by converting all text to lowercase and removing missing values, stop words, punctuation, web links, numbers, special characters and anything that is irrelevant and therefore can reduce the effectiveness of sentiment analysis. Some examples of irrelevant and uninformative words are shown in Table 2.[5]

Given that the comments were not labeled, we used a pre-trained model to automatically classify the texts into positive, negative or neutral sentiments. For this we used a BERT model like **nlptow/bert-base-multilingual- uncased-sentiment** , which is designed for this type of task and supports several languages. Then, the labels are used to train the LSTM model to recognize patterns in text sequences that correspond to these sentiments.

Table.3. Uninformative words in the dataset

Noises Types	Example
Stop Words	Svp, 5km,où,dans,lol, 💜, 😊,Humm,ok, ect.
Special Characters	!,¤, ?,@,&, », etc.
Duplicate Words	Belllle, Yessss, etc…

3.1.2 Model LSTM (Long Short-Term Memory)

LSTM is a more advanced type of Recurrent Neural Network that addresses the gradient explosion and vanishing gradient issues that are common in traditional RNNs. LSTM can handle sequences of varying lengths and efficiently process sequential data, which helps to mitigate the problem of information loss in recurrent neurons [5].

Several key components make up the Long Short-Term Memory (LSTM) architecture, including the input word (Xt) at the current time step, the cell state (Ct), the temporary cell state (ćt), the hidden layer state (Ht), the forgetting gate (Ft), the memory gate (It), and the output gate (Ot). The LSTM model has three main stages: forgetting, selective memory, and output. During the forgetting stage, the LSTM selectively forgets information stored in cells while retaining important information. This process is governed by the forgetting gate (Ft).

Following the forgetting stage, the LSTM selectively "remembers" information from the input cells. Important information is highlighted and retained, whereas unimportant information is ignored. The memory gate (Ct) is essential for controlling this stage. Finally, the output stage determines which information will be used for the final output. The output gate (Ot) controls this stage. The LSTM model works in a sequential manner, with each stage influencing and being influenced by the previous one.

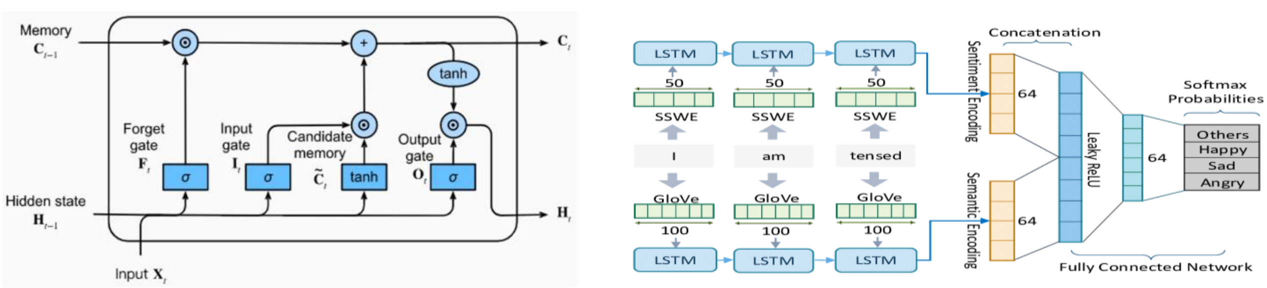

Fig4: Architecture des LSTM

Table.4. Sample of Content Based dataset

	date	postTitle	profileName	text	Cleaned_text	Sentiment_Label	Sentiment_Score
0	2020-11-02	Chutes de Tello \nSituées à 45km de Ngaoundéré...	Jane Elombe	Alfred	alfred	positive	0.346164
1	2020-09-01	Chutes de Tello \nSituées à 45km de Ngaoundéré...	Blanche Abossolo	Guy J.	guy	neutral	0.263435
2	2020-08-30	Chutes de Tello \nSituées à 45km de Ngaoundéré...	Emilie Dijon	Le pays est beau	pay beau	positive	0.405813
3	2020-08-30	Chutes de Tello \nSituées à 45km de Ngaoundéré...	Manuella Djedoua	Belel, l'endroit où mon père est née...	belel endroit où père née	neutral	0.256499
4	2020-08-30	Chutes de Tello \nSituées à 45km de Ngaoundéré...	Evedassi Dassi Eve	Ma question est la suivante,vous offrez des se...	question suivante offrez service guide tourist...	neutral	0.363380
5	2020-08-29	Chutes de Tello \nSituées à 45km de Ngaoundéré...	Faissal Adamou	A 5 km de mon village sadol Yaya (bera)	a km village sadol yaya bera	negative	0.278091
6	2020-08-29	Chutes de Tello \nSituées à 45km de Ngaoundéré...	Saraounia Saraounia	Mon beau village ❤	beau village	positive	0.431664

3.3 Geographical Position

Here, a division is made according to the 10 regions of Cameroon and the regions where the tourist destinations are represented (parks, reserves, caves, chiefdoms, museums etc.) are used. Geographic coordinates such as longitude and latitude are taken into account then we calculate the Euclidean distance between two geographical points using the Euclidean distance formula [8]. Then for a new user, we calculate the distance between their position and all the sites, then recommend the closest sites based on the scores.

$$d_{Eucl}(p,q) = \\ = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \ldots + (q_n - p_n)^2} = \\ = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}$$

Equation1: Euclidean distance Formula

sites touristiques	score
Parc national de Bouba Ndjida	0.86
Chutes de Velempeng	0.68
Rochers de Gashiga	0.59
Les gorges de Kola dans le Nord	0.56
Mine de kaolin de Balengou	0.40

Fig5: Example of scores for some sites

3.4 Hybrid system

After identifying the user's preferences, the system then applies an adjusted cosine similarity approach to find other destinations that might be of interest to a particular user. The formula for adjusted cosine similarity is given in equation2.[6]

$$sim(x,y) = \frac{\sum_{s \in S_{xy}} (r_{xs} - \bar{r_x})(r_{ys} - \bar{r_y})}{\sqrt{\sum_{s \in S_{xy}} (r_{xs} - \bar{r_x})^2 (r_{ys} - \bar{r_y})^2}}$$

Equation2: Cosine Similarity

- S_{xy} represents the set of images commonly rated by both users x and y.
- R_{ys} shows ratings provided by user *x* to book s .
- $\overline{R}$x shows the average score of rating user x provided to images

Once similarities between items are identified, a weighted sum approach is used to predict the ratings of destinations that are not yet rated by active users. The item-based CF system makes recommendations in two steps. First, they calculate the similarity between destinations that the user has liked in the past and those that the user has never experienced before. Second, they recommend and select the top 10 best destinations that meet the user's needs and match the previous user, adding the sentiment score and location score. Mathematically it can be represented as total recommended set.

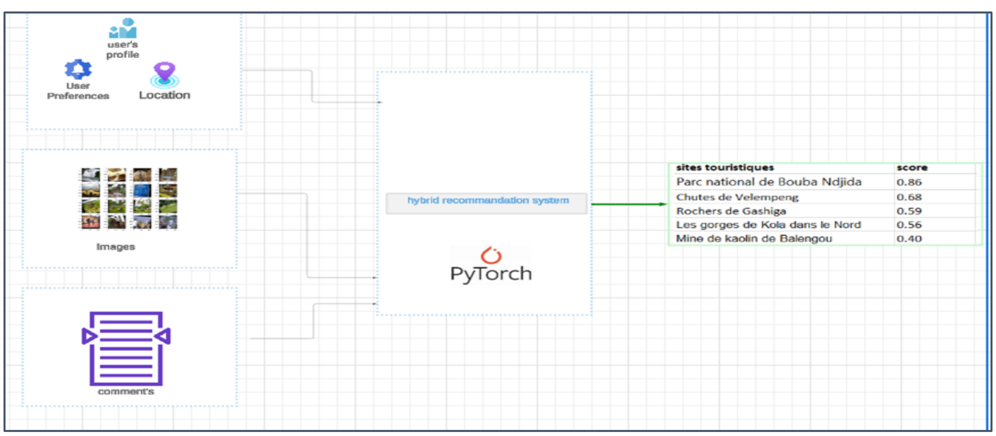

Fig6: Our Hybrid System Architecture

4 Results and discussions

To evaluate the performance of our system, we used a dataset of tourist destinations and corresponding user ratings. We split the dataset into training and test data, and we used the test set to evaluate the performance of our system using F $_1$-Score, precision@10, recall@10 as an evaluation metric. The below results from our analyzes show that the proposed hybrid system outperforms both collaborative filtering and content-based approaches and also takes into account the geographic area of a new user. In any recommendation system, the two essential blocks are the users and the elements (here the elements mean the destinations) to be recommended.

System	Precision@10	Recall@10	F_1-Score
Con-B	0.183	0.185	0.180
Col-F	0 415	0.396	0.411
Hybrid	**0.466**	**0.464**	**0.482**

Table.5.Résultats Métriques d'évaluation

5 Conclusions and future work

In this work, we proposed a hybrid recommendation system that uses a multimodal approach that combines the analysis of text, images and geographic areas to make recommendations to a user. The system is based on a pre-trained model to extract image characteristics as well as another to analyze user reviews. The scores are then calculated and added to obtain a score and produce the Top-10 recommendations likely to interest the user. Comparison of our model with techniques based solely on a single approach (content and collaborative) used by tourist sites in Cameroon shows a remarkable improvement.

Future work will involve **integrating more data** to enrich the recommendations and refining the deep learning techniques used to improve the accuracy and relevance of the recommendations. We are also thinking of exploring fuzzy logic or the age of users to evaluate the behavior of our system.

References

1. Lee, SW; Jiang, G.; Kong, HY; Liu, C. A difference of multimedia consumer's rating and review through sentiment analysis. Multimed. Tools Appl. 2021, 80, 34625–34642. [CrossRef]
2. Ebadi, A.; Krzyzak, A. A hybrid multi-criteria hotel recommender system using explicit and implicit feedback. Int. J. Comput. Inf. Eng. 2016, 10, 1450–1458.
3. Hybrid Recommendation System for Tourism using Deep Learning Mohamed BADOUCH, Faculty of Sciences, Ibn Zohr University, Agadir, Morocco Mehdi BOUTAOUNTE, National School of Commerce and Management, Ibn Zohr University, Dakhla, Morocco
4. Enhancing Collaborative Filtering-Based Recommender System Using Sentiment Analysis Ikram Karabila 1,*, Nossayba Darraz 1 , Anas El-Ansari 2 , Nabil Alami 3 and Mostafa El Mallahi 1
5. Amzad, H., Vijayalakshmi, K. : Système de recommandation touristique : une revue systématique. Revue internationale de recherche et de technologie en ingénierie (IJERT) Vol 10
6. TourOptiGuide: A Hybrid and Personalized Tourism Recommendation System: Arfaoui Nouha Research Team in Intelligent Machines (RTIM) Ejbali Ridha Research Team in Intelligent Machines (RTIM), Posted Date: May 8th, 2024
7. Airbnb recommendation system using Aspect-based sentiment analysis: hybrid approach: Manav Mishra STUDENT NUMBER: 2038836, THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN DATA SCIENCE & SOCIETY DEPARTMENT OF COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE SCHOOL OF HUMANITIES AND DIGITAL SCIENCES TILBURG UNIVERSITY, Tilburg University School of Humanities and Digital Sciences Department of Cognitive Science & Artificial Intelligence Tilburg, The Netherlands May 2021
8. I MPSI – Option Informatique Ann´ee 2001, Deuxi`eme TP Caml Vincent Simonet (http://cristal.inria.fr/~simonet/)