مجلة جامعة طرطوس للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (6) العدد(7) 2022
Tartous University Journal for Research and Scientific Studies - engineering Sciences Series Vol. (6) No. (7) 2022

تحسين نظام التوصية لدعم السّياحة في المنطقة العربية بالاعتماد على الشبكات الاجتماعية

د. ماهر إبراهيم*

رهف علي جديد**

(تاريخ الإيداع 4/20 /2022 . قُبِل للنشر في 10 /8/2022)

□ ملخّص□

شهدت الأعوام الأخيرة تقدّماً هائلاً في التكنولوجيا والاعتماد عليها في الكثير من المجالات من بينها السياحة، حيث شغلت حيّزاً كبيراً ضمن هذا المجال، هذا التّقدّم أدّى إلى مصادر لانهائيّة من المعلومات المتعلّقة بالسّياحة، والفيسبوك يعتبر مثال عن أحد هذه المصادر، وهو عبارة عن منصّة حيث يقوم النّاس بنشر خبرتهم وتقييماتهم بشكل يومي. وبذلك نجد العديد من تقنيّات التجميع والتّنبّؤ الجديدة التي تساعد المستخدمين على اختيار الأماكن المناسبة التي تلبّي احتياجاتهم و / أو تفضيلاتهم من عدد هائل من الخيارات المتاحة. تقدّم هذه الورقة نظام توصية بالأماكن السّياحيّة يقوم على معرفة نوع المناطق السّياحيّة التي يهتم بها السّيّاح، وإيجاد التّوصيات المناسبة التي تستند إلى البيانات التي يتم تجميعها ديناميكيّاً واستقراءها من بيانات تسجيل الوصول على فيسبوك. تمّ استخدام نظام التّوصية الهجين حيث يدمج بين مزايا أنظمة التّوصية المدركة للسياق من ناحية تجميع المستخدمين بناءً على اهتماماتهم ومزايا أنظمة التّوصية التّعاونية من ناحية إيجاد المستخدمين الذين لديهم اهتمامات مشابهة للمستخدم الحالي. يمكن للنظام الحالي أن يتكيّف مع سلوك المستخدم المتغيّر بشكل ديناميكي وبالتالي يعطي التّوصيات بما يناسب التغيير الحاصل. تمّ توضيح أهميّة البيانات المتاحة على الشبكات الاجتماعية من أجل تحسين جودة عمليّة التوصية بالأماكن السياحية التي تناسب اهتمامات المستخدمين.

الكلمات المفتاحيّة: السّياحة، أنظمة التّوصية، الشبكات الاجتماعيّة، تسجيلات الوصول، الخصوصيّة.

* مدرّس في قسم هندسة تكنولوجيا المعلومات – كلّيّة هندسة تكنولوجيا المعلومات والاتصالات – جامعة طرطوس – سوريا

** طالبة ماجستير في قسم هندسة تكنولوجيا الاتصالات – كلّيّة هندسة تكنولوجيا المعلومات والاتصالات – جامعة طرطوس – سوريا

مجلة جامعة طرطوس للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (6) العدد(7) 2022
Tartous University Journal for Research and Scientific Studies - engineering Sciences Series Vol. (6) No. (7) 2022

Improve the Recommendation system to support tourism in the Arab region depending on social media

Dr. Maher Ibrahim*
Rahaf Ali Jdeed**

(Received // 2022 . Accepted // 2022)

□ ABSTRACT

Recent years have witnessed enormous progress in technology and reliance on it in many fields, including tourism, where it occupied a large space within this field. This progress in the technology field has led to endless sources of information related to tourism, and Facebook is an example of one of these sources, a platform where people share their experiences and opinions on a daily basis. Thus, we find many new aggregation and prediction technologies that help users to choose the right places that meet their needs and/ or preferences from the huge amount of available data. This paper presents a system for recommending attraction places based on knowing what type of attractions tourists are interested in, and finding appropriate recommendations based on the data that is collected dynamically and extrapolated from the Facebook check-ins data. A hybrid recommendation system (RS) has been used which combines the advantages of both context-aware systems and collaborative systems. The current system can adapt the changing in user behavior dynamically and adapt the recommendations based on it, which improves the recommendation quality for the tourists.

Key Words: Tourism, Recommendation System (RS), Social Networks, Check-ins, personalisation.

* Teacher, Information Technology Engineering Department, Information and communication Technology Engineering, Tartous University, Syria.
** Master student, Information Technology Engineering Department, Information and communication Technology Engineering, Tartous University, Syria.

1. المقدّمة:

سجّلت صناعة تكنولوجيا المعلومات نموّاً هائلاً خلال العقد الماضي، ممّا أدّى بدوره إلى نمو كبير في حجم البيانات، بالتالي أصبح هناك حاجة لإدارة هذه البيانات. وتعدّ السّياحة من الأعمال التي تتطلّب معلومات مكثّفة، ومع ذلك يتوقّع السّيّاح المعاصرون الوصول إلى هذه المجموعة الكبيرة من المعلومات في أي وقت باستخدام الوسيلة المفضّلة لديهم. ونتيجة لذلك هنالك الكثير من المعلومات المتاحة، ممّا يؤدّي في بعض الأحيان إلى ما يُعرَف باسم الحمل الزّائد للمعلومات، وتعتبر هذه مشكلة حقيقيّة حيث يقوم السّائح في كثير من الأحيان بقراءة الكثير من المعلومات ويحتاج للكثير من الوقت للبحث عن المكان المناسب لاهتماماته ومكانه الجغرافي [1]. بالتالي تمّ التّوجه لاستخدام أنظمة تساعد المستخدم في إيجاد المكان الأكثر ملاءمة لاهتماماته تُعرّف بأنظمة التّوصية Recommendation System.

نظام التّوصية هو عبارة عن خوارزميّة تهدف لاقتراح بيانات ذات صلة بالمستخدم اعتماداً على تفضيلات سابقة أو تقييمات تمّ تقديمها من قبل المستخدم، أو تقترح بيانات ذات صلة للمستخدم اعتماداً على خصائص العنصر. ولهذه الأنظمة عدّة أنواع تنقسم بحسب تطبيقها [2] إلى: الأنظمة التّعاونيّة Collaborative systems، الأنظمة المُدركة للسياق Context-Aware systems، الأنظمة القائمة على المعرفة Knowledge-based systems والأنظمة الهجينة Hybrid systems. تقوم استراتيجيّة التّوصية بشكل عام على خطوتين أساسيّتين حيث تفهم هذه الأنظمة سلوك المستخدم من خلال بياناته التي تمّ الحصول عليها من الشبكات الاجتماعيّة Social Networks وتحليلها مع تفضيلات المستخدم، ثمّ إنشاء نقاط الاهتمام Points of Interests (POI) استناداً إلى الاهتمامات المستخرجة من الخطوة السّابقة [3].

قام I. Cenamor, et al. [4] بتقديم نظام يقوم بإنشاء خطط سياحيّة مخصّصة باستخدام المعلومات التي تم إنشاؤها من قبل المستخدم وجمعها من شبكة التواصل الاجتماعي minube، تمّ في هذه الدّراسة الاعتماد فقط على المعلومات المقدّمة من minube، على الرّغم من إمكانيّة الحصول على معلومات من Yelp وGoogle places.

وتمّ تقديم تصميم للحل الذي يمكن استخدامه لتحقيق مهمّة اختيار السّيّاح لموقع معيّن من حيث إمكانيّة الذهاب إليه بعد جمع المعلومات عنه من الشبكات الاجتماعية عند M. Ellouze, et al. [5]، يتم استخراج آراء السّيّاح من مصادر مختلفة وتحليلها لاستخراج كيف يتم النّظر إلى الفنادق من قبل عملائها من حيث الميّزات والأنشطة. حيث اقتصرت هذه الدراسة على آراء المستخدمين بالفنادق وتقييمهم لها.

اقترح R. Hassannia, et al. [2] وكالة برمجية للتوصية الذّكيّة تعتمد على الوكيل باستخدام منهج التّوصية الهجين والذي يأخذ التّواصل في الزمن الحقيقي في الحسبان. تمّ تصميم البرنامج استناداً إلى تقنيّة الوكيل وتعريف وتنفيذ 5 وكلاء عبر تطبيق الويب حيث يتم التواصل بين الوكلاء عن طريق بروتوكول contact net، بيّنت النتائج أن هذا التطبيق حسّن من معدّل التّوصية مقارنةً بالمنصّات الأخرى، ولكن لم يتم اختبار دراسة الحالة على العملاء الحقيقيّين، بالتالي يجب إجراء المزيد من المحاكاة والاختبار في البيئة العامّة.

2. أهميّة البحث وأهدافه:

تأتي أهمّيّة هذا البحث من كونه يقدّم مساهمة جديدة في عملية دعم القرار، حيث أنّ الكمّ الهائل من البيانات على الانترنت جعل من عمليّة اتّخاذ القرار حول الوجهة السّياحية التي تناسب متطلبات الأفراد واهتماماتهم عمليّة صعبة وتأخذ الكثير من الوقت في التّفكير، بالإضافة إلى قلّة أنظمة التّوصية الذّكية التي تدعم المنطقة العربيّة في مجال السّياحة وتعمل على تحليل معلومات المستخدمين من الشبكات الاجتماعية للحصول على اهتماماتهم واقتراح أماكن سياحية تناسبهم.

3. طرائق البحث وموادّه:

يقوم البحث على تطوير نظام يعمل على اقتراح الأماكن السّياحيّة المناسبة للمستخدمين بالاعتماد على تسجيلات الوصول Check-ins الخاصّة بهم التي يتم استخراجها من فيسبوك، حيث يتم اعتماد خوارزميّات التحليل لتحليل اهتمامات المستخدم وإيجاد المكان الأكثر ملاءمة له.

تمّ استخراج قاعدة البيانات الخاصّة بالأماكن السّياحيّة وتجميعها يدويّاً من خلال Facebook graph API وخرائط غوغل Google Map API.

قمنا باعتماد منهج تردّد العنصر Item Frequency من أجل الحصول على الفئات السّياحيّة الأكثر زيارة لمعرفة اهتمامات المستخدمين بكل فئة، وقياس درجة الاهتمام الخاصّة بالمستخدم وتمثيلها في نموذج خاص به user model.

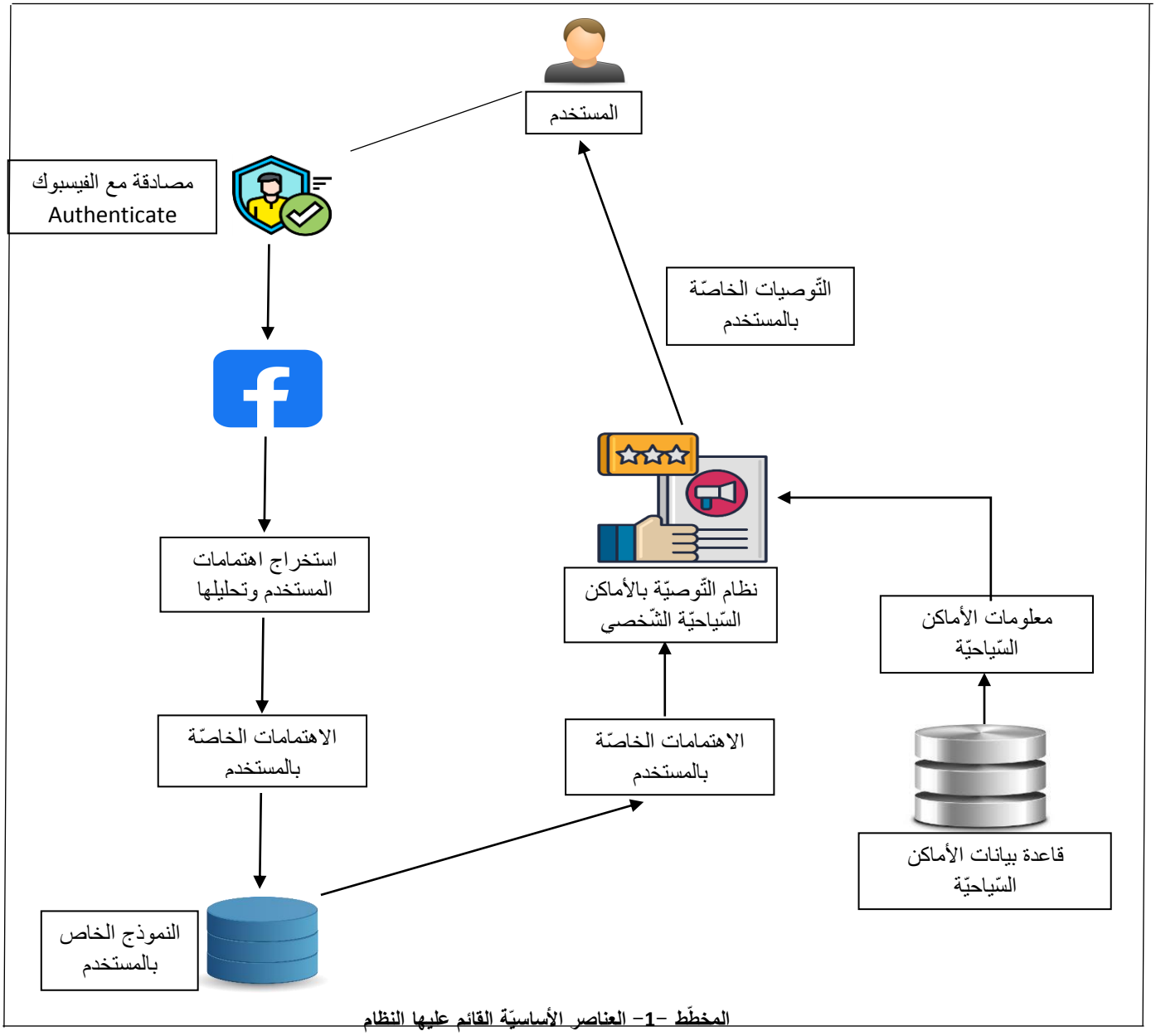
1.3 الخوارزمية المُقتَرحة:

تم في هذا البحث العمل على تطوير نظام يعمل على اقتراح الأماكن السياحيّة التي تناسب اهتمامات المستخدم والتي تمّ الحصول عليها عن طريق استخراج تسجيلات الوصول الخاصّة بالمستخدم على فيسبوك والعمل على تحليلها للحصول على الاهتمامات[6] .

قمنا بإضافة بارامتر خاص بالميزانيّة الخاصّة بالسّائح إلى نموذج تحليل الانحدار الخطّي المُطَبّق في عمليّة التّوصية بالأماكن السّياحيّة وهذا لم يتم التّطرّق إليه في الدراسات السّابقة، حيث يقوم المستخدم بإدخال الميزانية التي يريد استخدامها ويتم معالجتها في النظام للحصول على مكان يناسب هذه الميزانية.

يعتمد النّظام على ثلاثة عناصر رئيسيّة تمّ تلخيصها في المخطّط -1-:

- قواعد البيانات.
- استخراج اهتمامات المستخدم وتحليلها.
- عملية التّوصية بالمكان السّياحي المناسب.



المخطّط -1- العناصر الأساسيّة القائم عليها النظام

1.1.3 قواعد البيانات Datasets:

قمنا باستخدام قاعدتي بيانات، قاعدة بيانات خاصّة بالأماكن السّياحيّة وقاعدة بيانات خاصّة بمعلومات المستخدمين المُستّخرجة من الفيسبوك (سيتم الشرح عنها ضمن استخراج اهتمامات المستخدم).

- **قاعدة البيانات الخاصّة بالأماكن Attraction dataset:**

تمّ استخراج قاعدة البيانات وتجميعها يدوياً عن طريق الفيسبوك Facebook graph API وخرائط غوغل Google map API عن طريق برنامج خاصّ بهذه العمليّة Botsol Crawler.

تتألّف قاعدة البيانات الخاصّة بالأماكن من السّمات التّالية:

- اسم المكان Business Name.
- الصورة Image: وهي عبارة عن صورة لها علاقة بالمكان المذكور.
- اسم المحافظة التّابع لها State.

- العنوان Address.
- فئة المكان Category: إذا كان مطعم Restaurant، فندق Hotel، مكان أثري Archeological، أو مكان جذب سياحي Attraction.
- التقييم الخاص بالمكان الذي تمّ الحصول عليه عن طريق الفيسبوك Rating.
- عدد الإعجابات Likes.
- عدد تسجيلات الوصول Total check-ins.
- الموقع الجغرافي المؤلّف من Latitude, Longitude.
- عدد النجوم Stars.
- الوقت المناسب للمكان time weight.
- الميزانيّة الخاصّة بالمكان Budget.

- تم القيام **بمعالجة مُسبَقة لقاعدة البيانات Preprocessing** من أجل الحصول على بيانات نظيفة لاستخدامها في عملية التّحليل والتّوصية. حيث قمنا بحذف السّجلات المكرّرة للأماكن، معالجة الأخطاء الإملائية، معالجة السّجلات التي فيها نقص أو فراغ في أحد السّمات وتوحيد السّجلات.

2.1.3 استخراج وتحليل اهتمامات المستخدم:

تقوم هذه المرحلة باستخراج المعلومات الخاصّة بالمستخدم، وذلك باستخدام Facebook graph API. ومن أجل القيام بعملية استخراج البيانات الخاصّة بالمستخدم يجب أولاً على المستخدم أن يقوم بمصادقة حسابه ضمن النظام من أجل السّماح باسترجاع بياناته الخاصّة مثل تسجيلات الوصول، إعجاباته، تعليقاته وغيرها... يعمل بعدها النظام على بناء نموذج تفضيل للمستخدم بعد معالجة البيانات الخاصّة به التي تمّ استخراجها، ويتم تمثيل هذا النموذج ضمن بنية علائقيّة RDB.

في كل مرّة يقوم مستخدم جديد بالدّخول للنظام يتم إنشاء ملف خاص به ضمن هذه المرحلة بالنّسبة للأشخاص الذين ليس لديهم حساب على فيسبوك، سيتم اقتراح الأماكن لهم بحسب شعبيّة المكان الذي يتم قياسه بعدد مرّات تسجيل الدّخول إليه.

يُلخّص الجدول التّالي مثال عن البيانات المُستَخرَجة من الفيسبوك باستخدام Facebook graph API، ليتمّ استخدامها من أجل تحليل اهتمامات المستخدم في المرحلة التالية.

الجدول –1– بيانات تسجيل الوصول المُستَخرجة من فيسبوك لتُضاف إلى قاعدة البيانات

وصف البيانات	البيانات التي تمّ استخراجها
	{“place”: {
معرّف المكان	“id”:” 573491572765389”,
اسم المكان	“name”:” MountainBreezeResort1”,
موقع المكان	“location”: {
Latitude	“latitude”: 35.4316051243,
Longitude	“longitude”: 36.1410478181,
العنوان	“street”: “Kurdaha, Syria”}
رقم أو معرّف الرّسالة	}, “id”:”355502081315175”,
وقت تسجيل الوصول	“created time”: “2015-05-08T19:29:02+0000”}

1.2.1.3 مصفوفة اهتمامات المستخدم:

تُظهِر هذه المصفوفة درجة الاهتمام بفئات الأماكن السّياحيّة باستخدام بيانات تسجيل الوصول، حيث يُشير عدد تسجيلات الوصول في فئة معيّنة إلى درجة الاهتمام بالأماكن المتعلّقة بهذه الفئة.

قمنا بتطبيق منهج تردّد العنصر item frequency من أجل قياس درجة اهتمام المستخدم وتمثيلها ضمن نموذج المستخدم. يتم إسناد وزن لكلّ مكان سياحي في فئة ما بناءً على عدد مرّات تسجيل الوصول في ذلك المكان.

درجة الاهتمام l للمستخدم u تعطي درجة الاهتمام لكل فئة سياحيّة[6] :

$$l_u(c)= \{(c,l)|c \in C \text{ and } l \in L, l \text{ is in the range } [0,1]\}$$

l هي درجة الاهتمام التي تمّ إسنادها إلى الفئة c لمستخدم معيّن u. حيث تعبّر هذه الدّرجة عن نسبة عدد الأماكن السّياحيّة التي سجّل فيها مستخدم ما وصول ضمن فئة ما c من إجمالي عدد الأماكن السّياحيّة ضمن نفس الفئة، ويجب أن تكون ضمن المجال من 0 حتى 1.

C هي عدد العناصر ضمن الفئة c وهنا تمّ تجميعها في 4 فئات رئيسيّة مثلاً $l_u(c)$ = {(Archeological,0.50), (Restaurant,0.25), (Hotel,0.25)} حيث هذه الفئات هي مطاعم، فنادق، مناطق أثريّة ومناطق جذب سياحي. مجموع جميع الأوزان في $l_u(c)$ يجب أن يساوي الـ 1، بحيث يشير المجموع 0 إلى أن المستخدم غير مهتمّ بهذه الفئة والقيمة 1 تشير إلى درجة اهتمام عالية بالفئة.

2.2.1.3 تحليل اهتمام المستخدم:

يتم اختيار معلومات بشكل تجريبي تشير إلى مستخدمين لهم بيانات تسجيل وصول أكبر من عتبة معيّنة threshold وذلك من أجل إيجاد اهتمامات المستخدم.

لحساب درجة اهتمام المستخدم يتم تطبيق المعادلة التالية[6] :

$$I_u(c) = \frac{n_c}{\sum_{i=1}^{\wedge} n_i}, \quad 1 \leq c \leq \wedge \qquad (1)$$

حيث $l_u(c)$ هي درجة اهتمام المستخدم u بفئة معيّنة c، n_c هي عدد تسجيلات الوصول للفئة c. ويتم تسوية $l_u(c)$ لتصبح ضمن المجال من 0 إلى 1.

الخوارزمية التي تمّ استخدامها من أجل تحليل اهتمام المستخدم:

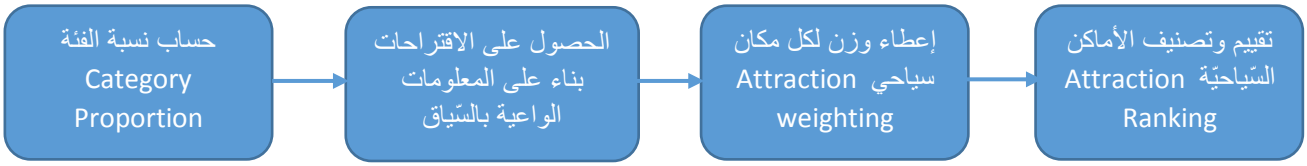
Input: All check-ins data of a user
Output: Interest level $I_u(c)$ in each category.
1: Remove duplicate check-ins from the same day
2: Set threshold = 10
3: if check-ins < threshold
4: Remove check-ins that are not attractions
5: Classify check-ins into four categories
6: Compute level of Interest $I_u(c)$ using Eq. (1)
7: Return $I_u(C)$.

3.2.1.3 التّوصية بالأماكن السّياحيّة:

هذه المرحلة من أجل اقتراح الأماكن السّياحيّة التي تناسب المستخدم بشكل كبير. حالما يتم معرفة اهتمامات المستخدم، يعمل النّظام على مطابقة اهتمامات المستخدم مع الأماكن السّياحيّة الموجودة اعتماداً على بعض الحسابات الرّياضيّة.

تُعطى الأماكن المناسبة تقييم rank بدرجات تنازليّة، ويُطبَّق المنهج من أجل اقتراح الأماكن السّياحيّة اعتماداً على تحليل تسجيلات وصول المستخدم عن طريق مطابقة مميّزات الأماكن السّياحيّة مع مميّزات المستخدم.

تتألّف هذه العمليّة من 4 مراحل أساسيّة:



- **حساب نسبة الفئة Category Proportion:**

يهتمّ النّاس عادةً بالعديد من الأشياء في الوقت نفسه، مثلاً يهتمّ المسافر عادةً بعدّة أماكن سياحيّة من فئات مختلفة. لذلك، لن يتم اقتراح الأماكن السّياحيّة في الفئة التي لها درجة اهتمام عالية فقط، بل سيتم اقتراح أماكن أخرى من فئات أخرى. بحيث تعمل هذه المرحلة على إيجاد النّسبة للأماكن في فئات مختلفة من أجل أن يتم اقتراحها للمستخدم حسب المعادلة التالية[6] :

$$R(c) = \frac{I_u(c)}{\sum_{k=1}^{4} I(k)}, \quad 1 \le c \le 4 \qquad \textbf{(2)}$$

R(c) هي نسبة الأماكن في فئة ما تم اقتراحها للمستخدم بناء على درجة الاهتمام التي تمّ حسابها في المعادلة 1 مقسومة على درجة الاهتمام لكلّ الفئات I(k).

- **الحصول على الاقتراحات بناء على المعلومات المدركة للسياق Context-awareness recommendations:**

حيث يعبّر السّياق عن أي معلومة يمكن أن يتم استخدامها من أجل أن تميّز الحالة الخاصّة بكيان ما. يمكن أن يكون السّياق في هذه الحالة أي معلومة تتعلّق بالحالة الخاصّة بالسّياحة مثل (الوقت، الموقع، الطقس وغيرها...)[7]. يمكن اعتبار الوقت من البيانات المهمّة والمفيدة في أنظمة التوصية، خاصّةً في مجال

السّياحة، حيث تتلقّى التّوصيات التي تدرك الوقت اهتمام أكبر بحيث أنّها تزيد من أداء التّوصية وتسهّل من عمليّة تتبّع تطوّر النّموذج الخاص بالمستخدم. بالتالي تمّ دمج المعلومات الخاصّة بالوقت في عمليّة التّوصية لتحسين دقّة نظام التوصية وذلك بتوفير هذه المعلومات حول الوقت الموصى به لزيارة المكان السّياحي (مثلاً: القيام باقتراح زيارة شواطئ البحر خلال موسم الشتاء لمسافر ما هو أمر خاطئ على الرّغم من أن تلك الشواطئ يمكن أن يكون لها درجات عالية من الاهتمام في ملف التعريف الخاصّ بالمسافر). بالتالي سيتم إعطاء توصيات للمستخدمين مناسبة لاهتماماتهم مع مراعاة للوقت الخاص بالزّيارة وبالتالي هذا سيولّد ثقة كبيرة جدّاً في توصيات النّظام.

بالتالي يمكن دمج أبعاد الوقت ضمن درجة الاهتمام لتعديل المعادلة 1، لتأخذ درجة الاهتمام معامل الوقت في الحسبان وتصبح كالتالي[6] :

$$I_u(c) = \{(c, l, t) | c \in C, l \text{ is in the range } [0,1], \text{ and } t \in T\}$$

حيث

t هي معلومة الوقت الهدف ويُشار إليه بالشهر المحدّد من قبل المستخدم.

يسمح استخدام معلومات السياق الزّمني للنّظام بتقديم توصيات متباينة اعتماداً على الوقت المستهدف من المستخدم. يؤثّر هذا العامل على نتائج التّصنيفات النّهائيّة. لذلك، يمكن أن يقترح النظام أماكن مختلفة للمستخدم إذا كان وقت الزّيارة المفضّل لديه مختلف.

من المعلومات السّياقيّة الأخرى التي يمكن الاعتماد عليها هي الميزانية الخاصّة بالسّائح من أجل الرّحلة التي يريد القيام بها، بالتالي تمّ إضافة المعلومات الخاصّة بميزانية كل مكان سياحي في عمليّة التّوصية من أجل زيادة دقّة نظام التّوصية وذلك بتوفير المعلومات الخاصّة بالميزانية المناسبة لكل مكان فمثلاً يمكن أن يكون السّائح من الطّبقة المتوسّطة ولا يستطيع الذّهاب إلى أماكن ذات ميزانيّة عالية جداً، بهذه الحالة سيتم اقتراح أماكن مناسبة تراعي اهتمامات المستخدم ووضعه الاجتماعي المالي وهذا سيعطي دقّة أعلى في التّوصيات[8]. وتصبح المعادلة السّابقة كما يلي:

$$I_u(c) = \{(c,l,t,b) | c \in C, l \text{ is in the range } [0,1], \text{ and } t \in T, b \in B\}$$

- **إعطاء وزن للمكان السّياحي Attraction weighting:**

يتم في هذه المرحلة إسناد وزن لكل مكان سياحي في كل فئة وذلك من أجل اختيار الأماكن ذات التّقييم R(c) الأعلى من أجل المستخدم[15] .

تعتمد عمليّة التّوزين على ثلاث متغيّرات هنا:

- **شعبيّة المكان Popularity of place (P):**

يرغب السّيّاح عادةً بزيارة الأماكن الأكثر شعبيّة أو شهرة في الوجهة السّياحيّة المختارة. تمّ الحصول على هذه المعلومات عن طريق استخراج عدد تسجيلات الوصول من الفيسبوك وخرائط غوغل. كلّما زاد عدد تسجيلات الوصول في الوجهة كلّما زادت شعبيّتها.

- **الوقت المناسب لزيارة المكان السّياحي Appropriate time to visit attraction (T):**

ناقشنا ضمن المعلومات المدركة للسياق أن المعلومات الخاصّة بالوقت هي شرط مهم لاتخاذ قرار بالسّفر في مجال السّياحة.

- **الميزانية المناسبة للمكان السّياحي Appropriate budget for the attraction (B):**

كما تم الحديث في المعلومات المدركة للسياق أن المعلومات الخاصّة بالميزانية المناسبة للمكان السياحي لها أثر كبير على اقتراح المكان المناسب للسائح. فبناءً على هذه الأسباب، يستخدم توزين المكان السّياحي نموذج تحليل الانحدار الخطّي لمراعاة جميع المتغيّرات من أجل التّوصيات السّياحيّة كما في المعادلة التالية:

$$W(p) = \alpha P(p) + \beta B(p) + \gamma T(p) \qquad (3)$$

حيث W(p) هي وزن المكان السّياحي، α هي الوزن الذي يُعطى لشعبيّة المكان، β هي الوزن المُعطى لميزانيّة المكان السياحي، γ هي الوزن الذي يُعطى للوقت المناسب لزيارة الأماكن السّياحيّة.

تُقاس شعبيّة المكان P(p) عن طريق عدد الإعجابات وتسجيلات الوصول للمكان. $N_{ch,p}$ عدد تسجيلات الوصول في المكان و $n_{li,p}$ عدد الإعجابات على فيسبوك للمكان. Max هي أعلى عدد لتسجيلات الوصول لكل مكان سياحي. تمّ تسوية المجال للمتغيّرات (الإعجابات وتسجيلات الوصول) لتأكيد أنّ البيانات لن يتم تحميلها بشكل زائد من قبل بعضها من ناحية مقاييس المسافة كما المعادلة[6] :

$$P(p) = \frac{n_{ch,p}}{\max(n_{ch,p})} + \frac{n_{li,p}}{\max(n_{li,p})} \qquad (4)$$

غالبا ذلك يساعد بارامتر الوقت على توصية المكان الصحيح في الوقت المناسب له. يوجد مجموعة من الشروط يجب تعريفها من أجل قيم الوقت التي تتراوح ضمن المجال [0,1] كالتالي:

- T(p)=0 تُشير إلى أنّ الوقت المناسب لزيارة المكان السّياحي هو فصل الشتاء. كلّما زادت القيمة كلما أصبح الوقت يقترب من فصلي الخريف والرّبيع.
- T(p)=0.4 تُشير إلى أنّ الوقت المناسب للزيارة أقرب لفصل الشتاء.
- T(p)=0.5 تُشير إلى أنّ الوقت المناسب لزيارة المكان السّياحي هو جميع الفصول بحيث يكون المكان صيفي شتوي.
- T(p)=0.7 تُشير إلى أن الوقت المناسب أقرب لفصل الصّيف.
- T(p)=1 تُشير إلى أنّ الوقت المناسب للزيارة أفضل في فصل الصّيف.

بالإضافة إلى أن كل مكان سياحي له ميزانيّة خاصّة والتي تعتمد على الخدمات المقدّمة وعدد النجوم وأهميّة المكان، لذلك يساعد هذا البارامتر على إعطاء التوصية المناسبة لميزانية المستخدم المُدخَلة.

كل قاعدة يتم تطبيقها على المعادلة (3) تعتمد على الوقت المُختار من قبل المستخدم وعلى الميزانيّة التي تمّ اختيارها.

➢ **تقييم وتصنيف الأماكن السّياحيّة Attraction Ranking:**

تقوم هذه المرحلة بقياس درجة ملاءمة الأماكن السّياحيّة لاهتمامات المستخدم المُعرّفة في الملف الخاص به. يتم تصنيف الأماكن السّياحيّة التي تمّ اقتراحها بترتيب تنازلي وتقديم قائمة بأفضل الأماكن السّياحيّة التي قد يفضّلها المستخدم. يُحسَب التّصنيف بناءً على المعادلة رقم 5-- [6]. إذا كان التّصنيف (p1) ذو درجة أكبر من التّصنيف (p2)، فهذا يُشير إلى أنّ p1 أكثر ملاءمة من p2 للمستخدم المحدّد.

$$Rank(p) = \frac{R(c) \times W(p)}{N} \qquad (5)$$

حيث N هو عدد التّوصيات بالأماكن السّياحيّة الموجودة.

2.3 تقييم الأداء ومناقشة النتائج:

تمّ القيام بعدّة تجارب من أجل تقييم النظام باستخدام قاعدة بيانات مستخرجة مؤلّفة من 40 مشارك بحسابات فيسبوك نشطة. جاء المشاركون من جامعة طرطوس، سوريا، تتراوح أعمار المشاركين بين 18 و35 عاماً، حيث كبار السّن أقل خبرة في استخدام تكنولوجيا المعلومات والقيام بعمليات تسجيل وصول خاصّة بهم. تمّ القيام بدعوة المشاركين عن طريق دعوة.

قمنا بتطبيق قاعدة البيانات الخاصّة بنا على النّظام السّابق ومن ثم تطبيقها على النّظام الخاص بنا ومقارنة النتائج التي تمّ الحصول عليها، وذلك لتقييم كفاءة النظام المقترح، الغرض من هذا التقييم هو دراسة فعاليّة إضافة الميزانيّة الخاصّة بالمستخدم للتوصيات إضافةً إلى تسجيلات الوصول الخاصّة به على فيسبوك مقارنةً مع النظام السابق، نظراً لأنه يجب اختيار أفضل n منطقة والتّوصية بها من بين جميع المناطق الموجودة، فهذا يؤثر في النهاية على دقّة التوصية ورضا المستخدم في استخدام النظام، حيث يبيّن الجدول التالي –2– اختلاف نظامنا عن النظام السابق.

الجدول –2– التحسين الخاص بنا

	النظام الخاص بنا	النظام السابق
استخدام سياق الوقت	✓	✓
استخدام الميزانية الخاصّة بالمستخدم	✓	X

لم نقم بمقارنة النتائج التي حصلنا عليها مع نظم التوصية الأخرى غير الذي قمنا بالاستناد عليه، لأن هذه التجارب تم إجراؤها على مجموعات مستخدمين مختلفة لهم بياناتهم وسماتها المختلفة. لذلك من الصّعب مقارنة نتائجنا مع نتائج الأنظمة الأخرى الموجودة. ولكن قمنا بمقارنة النتائج نظرياً، حيث أجريت التجارب لمقارنة جودة التوصية لطريقة التوصية التي قمنا باستخدامها مقارنة بالتطرق التقليدية القائمة على المحتوى والأنظمة التعاوني المنشورة. تمّ

استخدام المقاييس التقليدية المعياريّة للمقارنة من حيث معدّل الدّقّة (والتي تعبّر عن نسبة عدد التوصيات الإيجابية إلى عدد التوصيات الكلّيّة الإيجابية والسّلبية)، ودرجة التّصنيف ranking الناتجة من اجل تقييم كفاءة النظام[9] .

توفّر الدّقّة تقييماً مباشراً لمدى دقّة التّوصية [10]، لحساب الدقّة لكل مستخدم لدينا u، تمّ إعطاء درجة للمناطق وترتيبها وفقاً للمعادلة (5). يتم تعريف أنّ المكان السياحي مناسب أو ذو صلة للمستخدم إذا وجده جذّاباً. وبالتالي، يمكن تعريف الدّقّة على أنها نسبة المناطق السياحيّة المناسبة الموجودة في أعلى قائمة الأماكن المرتّبة في القائمة الناتجة عن المعادلة (5) ranking list، والمُشار إليها بـ N(n,u) من جميع الأماكن المُوصى بها. في الدراسة التي قمنا بها تمّ إعطاء N(u) قيمة 10 لأن النّظام سيقوم بتوصية أفضل 10 أماكن مناسبة للمستخدمين. يمكن تجميع الدّقّة من جميع المستخدمين للحصول على متوسّط الدّقّة لمجموعة بيانات الاختبار، ومن ثمّ يتم إعطاء متوسّط الدّقّة بواسطة (6)- [6]:

$$H(n,u) = \frac{\sum_u N(n,u)}{\sum_u N(u)} \qquad \textbf{(6)}$$

من الأهمية ملاحظة أنّنا لا نقوم بتقييم أداء النظام باستخدام recall لأنه يتم حسابها من النسبة N(n,u) من جميع الأماكن ذات الصّلة في قاعدة البيانات والمُشار إليها بـ R(n,u)، وهذا يعني أن قيمة recall تكون دائماً صغيرة لأن R(n,u) عادةً ما تكون أكبر بكثير من N(n,u) والتي يتم إصلاحها عند 10 (والتي تعبّر عن أعلى 10 توصيات).

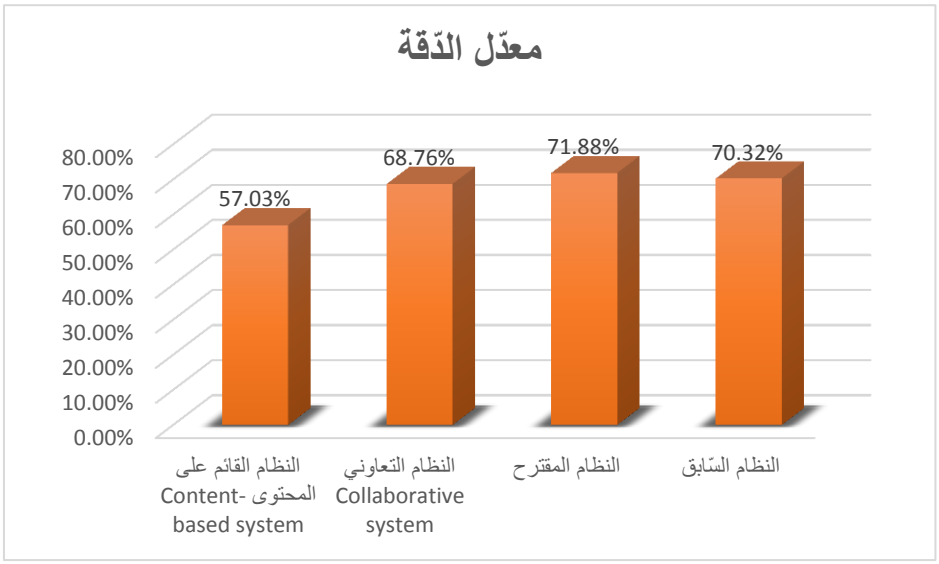
الجدول 3 مع المخطط –2– يوضّح لنا معدّل الدّقّة للنظام المقترح مقارنةً مع النظام القائم على النّهج التعاوني والنّظام القائم على المحتوى [6]. حيث نجد أنّ النظام المقترح أعلى معدّل من الأنظمة الأخرى وهذا يمكن شرحه بحقيقة أنّ النظام المقترح يستخدم بيانات تسجيلات الوصول الخاصّة بكل مستخدم والتي تؤثّر بشكل فعّال على تفضيلاتهم، حيث وصلت أعلى دقّة خاصّة بالنظام إلى 80.43% [6]، بالتالي هذا يساعد على التغلّب على مشكلة البداية الباردة، والتي هي سبب رئيسي في ضعف الأداء في النظام القائم على المحتوى [6]:

الجدول –3– معدّل دقّة النظام

اسم النظام	معدّل الدّقة
النظام القائم على المحتوى Content-based system	57.03%
النظام التعاوني Collaborative system	68.76%
النظام المقترح	71.88%
النظام السّابق	70.32%

من أجل التحقّق في الاختلاف الإحصائي بين هذه الطرق، تمت مقارنة متوسّط الدّقة لجميع الأساليب السابقة ضمن برنامج SPSS الإحصائي باستخدام ANOVA أحادي الاتجاه، ومن ثمّ طريقة Bonferroni

لمقارنات متعدّدة [12] [11] كما في الجدول 4. توضّح النتيجة أنهم مختلفون بشكل كبير حيث (F=13.48, p-value<0.001).



2- يوضّح معدّل الدّقّة الأنظمة

المخطط – الفرق في بين المذكورة

الجدول –4– مقارنة إحصائية بين الطرق المذكورة

اسم النظام	n	$\overline{x}$	S.D	F	p-Value
النظام القائم على المحتوى Content-based system	**10**	**0.57**[a]	**0.08**	**13.2**	**<0.001**
النظام التعاوني Collaborative system	10	0.69b	0.04		
النظام المقترح	10	72c	0.06		
النظام السّابق	10	70d	0.06		

ملاحظة: a,b,c,d تشير إلى فروق ذات دلالة إحصائية

حيث n: تعبّر عن عدد التوصيات التي تمّ الحصول عليها، x: معدّل الدّقّة الخاص بكل نظام، S.D: Standard Deviation الانحراف المعياري، F: تعبّر عن نسبة التّباين بين المتوسطات الحسابية الخاصّة بالعيّنة / التّباين ضمن العيّنة نفسها، حيث كلّما زادت قيمة F ضمن حسابات ANOVA كلما كان التّباين أعلى[11] ، P-value: هي احتمال يقيس الدليل ضد فرضيّة العدم.

هذه الدراسة تؤكّد أن بيانات تسجيل الوصول على فيسبوك يمكن أن تكون مفيدة وتستخدم لتحليلات وتوصيات بتفضيلات شخصية.

4. الأفكار المستقبليّة:

النظام الخاص بنا يعتمد فقط على بيانات تسجيل الوصول على فيسبوك الخاصة بالمستخدم، سيتم العمل على إضافة اهتمامات ومشاعر المستخدم من ناحية إعجاباته على فيسبوك وتحليل بياناته للعمل على الحصول على توصيات أفضل ومناسبة أكثر لشخصيّة المستخدم واهتماماته، حيث تلعب الإعجابات الخاصّة بالمستخدم على فيسبوك دوراً كبيراً في تحديد مشاعره تجاه مكان أو شيء معيّن[16]، بالتالي ستزيد دقّة التّوصية عند استخدام هذه الحالة.

5. المراجع:

[1] Meehan, K., Lunney, T., Curran, K., & McCaughey, A. (2013, March). Context-aware intelligent recommendation system for tourism. In *2013 IEEE international conference on pervasive computing and communications workshops (PERCOM workshops)* (pp. 328-331). IEEE.

[2] Hassannia, R., Vatankhah Barenji, A., Li, Z., & Alipour, H. (2019). Web-based recommendation system for smart tourism: Multiagent technology. *Sustainability*, *11*(2), 323.

[3] Menk, A., Sebastia, L., & Ferreira, R. (2019). Recommendation systems for tourism based on social networks: A survey. *arXiv preprint arXiv:1903.12099*.

[4] Cenamor, I., de la Rosa, T., Núñez, S., & Borrajo, D. (2017). Planning for tourism routes using social networks. *Expert Systems with Applications*, *69*, 1-9.

[5] Ellouze, M., Turki, S., Djaghloul, Y., & Foulonneau, M. (2017, September). Social Networks Based Framework for Recommending Touristic Locations. In *International Conference on Computational Collective Intelligence* (pp. 172-181). Springer, Cham.

[6] Kesorn, K., Juraphanthong, W., & Salaiwarakul, A. (2017). Personalized attraction recommendation system for tourists through check-in data. *IEEE access*, *5*, 26703-26721.

[7] Christensen, I., Schiaffino, S., & Armentano, M. (2016). Social group recommendation in the tourism domain. *Journal of intelligent information systems*, *47*(2), 209-231.

[8] Bin, C., Gu, T., Sun, Y., & Chang, L. (2019). A personalized POI route recommendation system based on heterogeneous tourism data and sequential pattern mining. Multimedia Tools and Applications, 78(24), 35135-35156.

[9] Deutschman, A. (2022, March 24). Recommender Systems: Machine Learning Metrics and Business Metrics.
https://neptune.ai/blog/recommender-systems-metrics

[10] Steck, H. (2010, July). Training and testing of recommender systems on data missing not at random. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 713-722).

[11] Weisstien, Eric W. "ANOVA." From MathWorld—A Wolfram Web Resource. https://mathworld.wolfram.com/ANOVA.html

[12] Weisstein, Eric W. "Bonferroni Correction." From MathWorld--A Wolfram Web Resource. https://mathworld.wolfram.com/BonferroniCorrection.html

[13] Manjare, P. A., Vninawe, M. P., Dabhire, M. M., Bonde, M. R., & Charhate, M. D. (2016). Recommendation System Based on Tourist Attraction. International Research Journal of Engineering and Technology, 3(04).

[14] Mr. Maher Ibrahim & Dr. S.A.M. Rizvi, A hybrid Approach for Better Utilization of Metadata Schema Registries, first national conference on next generation computing & information systems, CD proceedings under knowledge and Data Engineering section, S. No: 9, 12-13 May 2007, Model Institute of Engineering and Technology (MIET), Jammu, Kotbhalwal, India.

[15] Mr. Maher Ibrahim & Dr. S.A.M. Rizvi, A Framework for Building Standard System Data Dictionary, proceedings of the first international conference on emerging technologies and applications in engineering technology and sciences (ICETAETS), volume 3, paper number 3, Computer Science Department, Saurashtra University, 13-14 January,2008. Rajkot, Gujarat, India

[16] Dr. S.A.M. Rizvi & Mr. Maher Ibrahim, The Chart Representation of the System Data Dictionary, invertis international journal for science and technology, 2008