الأكاديمية العربية الدولية
Arab International Academy

الأكاديمية العربية الدولية

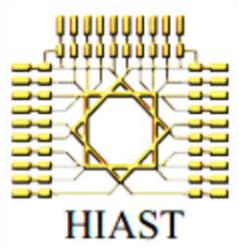
المقررات الجامعية

www.aiacademy.info | care@aiacademy.info

الجمهورية العربية السورية

المعهد العالي للعلوم التطبيقية والتكنولوجيا

قسم النظم المعلوماتية



ماجستير في المعلوماتية

بناء نظام توصية متعدد المعايير

خوارزمية سَلسَلة المعايير Criteria Chaining

أعدت هذه الأطروحة لنيل درجة

الماجستير في النظم المعلوماتية ودعم القرار

إعداد

م. ابراهيم الابراهيم

إشراف

د. ياسر رحال

كانون الأول 2018

المعهد العالي للعلوم التطبيقية والتكنولوجيا
Higher Institute for Applied Sciences & Technology

المعهد العالي للعلوم التطبيقية والتكنولوجيا مؤسسة حكومية للتعليم العالي أحدثت بموجب المرسوم التشريعي رقم /٢٤/ لعام ١٩٨٣، وذلك بهدف إعداد كوادر علمية متميزة من مهندسين وباحثين للإسهام الفاعل في عملية التطوير العلمي والتنمية في الجمهورية العربية السورية.

يمنح المعهد العالي درجة الإجازة في الهندسة في الاتصالات والمعلوماتية والنظم الإلكترونية والميكاترونيكس وعلوم وهندسة المواد وهندسة الطيران. يقبل المعهد العالي لدراسة هذه الاختصاصات شريحة منتقاة من المتفوقين في الشهادة الثانوية من الفرع العلمي. يتيح المعهد العالي أيضاً برامج ماجستير أكاديمي في نظم الاتصالات وفي التحكم والروبوتيك وفي نظم المعطيات الكبيرة ونظم المعلومات ودعم القرار وفي علوم وهندسة المواد وعلوم وهندسة البصريات. وأخيراً، يمنح المعهد العالي درجة الدكتوراه في الاتصالات والمعلوماتية ونظم التحكم والفيزياء التطبيقية. تُحدث في المعهد العالي اختصاصات جديدة بحسب متطلبات سوق العمل وتوجهات البحث والتطوير المحلية والعالمية.

يمارس المعهد العالي عبر جهود أطره وفعالياته العلمية المختلفة نشاطاً حثيثاً في البحث والتطوير، إذ ينفذ مشاريع ودراسات واستشارات متنوعة لصالح الجهات العامة والخاصة في القطر، كما يتعاون مع جهات خارج القطر في بعض المشاريع البحثية والتطويرية. يسعى المعهد أيضاً، عبر دورات تدريبية نظرية وعملية متاحة للقطاعين العام والخاص وللأفراد، إلى إفادة أوسع فئة من المهتمين من إمكانيات أطره العلمية ومختبراته. واستكمالاً لدوره الرائد في مجال التعليم ونشر العلم، ينشر المعهد العالي كتباً علمية عالية المستوى من نتاج أطره، منها ما هو تدريسي يوافق المناهج في المعهد العالي ويفيد شريحة واسعة من الطلاب الجامعيين عموماً، ومنها ما هو علمي ثقافي. يتيح المعهد العالي بعضاً من منشوراته على موقعه على الشابكة، كما يتيح إمكانية الاطلاع على رسائل الماجستير والدكتوراه المنفذة في المعهد العالي وعلى بعض منشورات طلابه وأطره من المقالات العلمية.

المعهد العالي للعلوم التطبيقية والتكنولوجيا، الجمهورية العربية السورية، دمشق، ص.ب ٣١٩٨٣
Higher Institute for Applied Sciences & Technology – HIAST
P. O. Box 31983, Damascus, Syrian Arab Republic
هاتف ٠٠٩٦٣١١٥١٢٣٨١٩ - فاكس ٠٠٩٦٣١١٥١٤٠٧٦١
بريد إلكتروني contact@hiast.edu.sy
موقع إلكتروني www.hiast.edu.sy

شكر

أتوجه بخالص الشكر والتقدير للدكتور ياسر رحال الذي لم يبخل بإشرافه وإرشاده في هذه العمل.

كما أتوجه بالشكر لكافة أعضاء الكادر التدريسي والإداري لماجستير نظم دعم القرار لما قدموه خلال فترة الماجستير من جهد ومعرفة تتجاوز المعارف المباشرة التي تلقيتها خلال دراستي في هذا الماجستير.

لا يمكن أن أنسى شكر المعهد العالي للعلوم التطبيقية والتكنولوجيا لإتاحته الفرصة لي لدراسة هذا الماجستير والحصول على شهادته كوني أحد طلابه السابقين ضمن مرحلة الإجازة الجامعية ومهندسيه خلال فترة عملي السابقة حتى الآن.

إهداء ...

إننا محكومون بالأمل... إليكِ سورية...

إلى من لم يبخلوا من أجل دفعي إلى طريق النجاح... والدي ووالدتي

إلى من كان حافزي إلى دخول المعهد العالي للعلوم التطبيقية والتكنولوجيا... أخي نديم

إلى من كان سندي في أصعب الظروف ورفيقي في أحلى اللحظات... زوجتي نجود

إلى شمعتي في الحياة... ابنتي ألفا

جدول المحتويات

1 الملخص v

2 الفصل الأول: مدخل لموضوع البحث وأهميته 7

1.2 تمهيد 8

2.2 دوافع البحث 9

3.2 إشكالية البحث 9

4.2 فكرة الحل المقترح 10

3 الفصل الثاني: الإطار النظري العام 11

1.3 مقدمة 12

2.3 أنظمة التوصية الرقمية التقليدية وحيدة المعيار 13

3.3 أنظمة التّوصية القائمة على المحتوى 14

4.3 أنظمة التّوصية القائمة على مبدأ التّرشيح التّعاونيّ Collaborative Filtering 15

5.3 التوصية الرقمية كمسألة صنع قرار متعددة المعايير 17

1.5.3 غرض القرار 18

2.5.3 عائلة المعايير 18

3.5.3 نموذج التفضيل العام 19

4.5.3 إجراء دعم القرار 20

6.3 إطار نظم دعم القرار متعددة المعايير في نظم التوصية الرقمية 21

1.6.3 البحث والتصفية في المحتوى متعدد الواصفات ونمذجة التفضيلات 22

2.6.3 استراتيجيات التوصية متعددة الأهداف 22

3.6.3 التفضيل القائم على التقييم متعدد المعايير 22

7.3 الخلاصة 23

4 الفصل الثالث: الدراسة المرجعية 24

1.4 نظم التوصية متعددة المعايير Multi-Criteria Rating Recommendation 25

1.1.4 التوصية وحيدة المعيار التقليدية 25

2.1.4 توسعة نظم التوصية التقليدية لتشمل التقييم متعدد المعايير 26

2.4 تضمين التقييم متعدد المعايير أثناء التوقع Engaging Multi-Criteria Ratings During Prediction 30
1.2.4 المقاربات التجريبية Heuristic Approaches 30
2.2.4 المقاربات المعتمدة على النموذج Model-based Approaches 36
3.4 تضمين التقييم متعدد المعايير أثناء التوصية Engaging Multi-Criteria Ratings During Recommendation 38
4.4 أمثلة تعدد المعايير Multi-Criteria Optimization 38
1.4.4 تصميم ترتيب كلي لتوصية المواد 39
2.4.4 إيجاد التوصيات الأمثلية من نمط Pareto 40
3.4.4 استخدام التقييم متعدد المعايير كمعاملات تصفية للاقتراحات 42
5.4 مناقشة الأعمال المستقبلية 43
1.5.4 تطوير مقاربات جديدة للتقييمات متعددة المعايير 43
2.5.4 توسعة التقنيات الحالية للنظم متعددة المعايير 45
6.4 محرّكات توصية مشابهة 45
1.6.4 Recombee 45
2.6.4 Suggest Grid 46
3.6.4 خوارزمية Yong Z. 47
7.4 خلاصة 47
5 الفصل الرابع: تصميم الحل المقترح والنتائج 48
1.5 مقدمة 49
2.5 تذكرة بإشكالية البحث 49
3.5 بناء عائلة المعايير 50
4.5 مجموعة البيانات المستعملة Trip Advisor Dataset 51
5.5 الخوارزمية المقترحة: سلسلة المعايير Criteria Chaining 52
6.5 تطبيق خوارزمية سلسلة المعايير على بيانات موقع Trip Advisor 56
1.6.5 مقارنة النموذج المقترح مع النموذج التقليدي 57
2.6.5 تأثير اختيار قيمة حجم عينة الجوار على أداء نظام التوصية 58
3.6.5 مقارنة الخوارزمية المقترحة مع خوارزمية Criteria Chains 58
7.5 المخطط الداخلي لعمل النظام البرمجي 59

8.5 بيئة العمل البرمجية 61

6 الفصل الخامس: الخاتمة والآفاق المستقبلية..... 62

1.6 الخاتمة 63

2.6 الآفاق المستقبلية..... 64

7 الملاحق..... 65

1.7 ملحق 1: أنظمة التوصية وحيدة المعيار 66

1.1.7 التّوصية باعتماد طريقة الجوار 66

2.1.7 مميزات اعتماد طريقة الجوار Neighborhood Approach..... 66

3.1.7 التنبّؤ بالتّقييم اعتماداً على المستخدم 67

4.1.7 حساب التّرابط..... 69

5.1.7 مصداقية أوزان التّرابط 72

6.1.7 اختيار الجوار 73

7.1.7 عنقدة المستخدمين والأغراض 74

8.1.7 اختبار النّظام Testing..... 77

8 المراجع..... 79

قائمة الأشكال

الشكل 1 نظام توصية أفلام وحيد المعيار 26
الشكل 2 نظام توصية أفلام متعدد المعايير (التقييمات لكل فيلم: التقييم الكلي، القصة، الحركات، الإخراج، المؤثرات البصرية)..... 27
الشكل 3 مقاربة التابع التجميعي (مثال على 3 معايير إفرادية) 37
الشكل 4 مثال على النقاط المسيطرة (المطاعم الأفضل) في فضاء ثنائي الأبعاد 41
الشكل 5 مثال يوضح حالة استخدام لخدمة محرك التوصيّة Recombee..... 46
الشكل 6 مثال يوضح حالة استخدام لخدمة محرك التوصيّة SuggestGrid 47
الشكل 5 مثال عن التقييمات متعددة المعايير في موقع Trip Advisor..... 51
الشكل 6 خوارزمية سلسلة المعايير 55
الشكل 7 بنية المعطيات الممثلة لتقييمات بيانات Trip Advisor..... 56
الشكل 8 تناقص الخطأ في عمليّة التنبؤ مع ازدياد عدد الجوار. 58
الشكل 9 المخطط الداخلي لعمل النظام..... 59
الشكل 10 مرحل تقسيم المعطيات في تقنية Cross-Validation..... 77

قائمة المصطلحات والاختصارات

المصطلح أو الاختصار	المعنى باللغة الإنكليزية	المعنى باللغة العربية
MCRS	Multi-Criteria Recommendation Systems	نظم التوصية متعددة المعايير
CF Recommenders	Collaborative Filtering Recommenders	محركات التوصية المعتمدة على الترشيح التعاوني
CB Recommenders	Content Based Recommenders	محركات التوصية المعتمدة على المحتوى
Decision Support		دعم القرار
Heuristic		تجريبي
Criteria		معايير
SVR	Support Vector Regression	توابع التجميع المعقدة
Linear Regression		الانحدار الخطي
MAE	Minimum Absolute Error	الخطأ التربيعي الأصغري بالقيمة المطلقة
CV	Cross Validation	التحقق المزدوج

الملخص

أصبحت أنظمة التوصية، أو كما تعرف أيضاً بمحركات التوصية، مجال بحث مهم يتم تطبيقه في العديد من المجالات. بالإضافة إلى ذلك، جرى تطوير التقنيات خلف هذه الأنظمة خلال الزمن. في العموم، تساعد هذه الأنظمة المستخدمين على إيجاد المنتجات أو الخدمات (مثل الكتب أو الموسيقى) المطلوبة من قبلهم، وذلك عن طريق تحليل وتجميع نشاطات المستخدمين الآخرين، عن طريق المراجعات Reviews بشكل أساسي، ومن ثم تشكيل التوصيات. تساعد التوصيات على تسهيل إجراء صنع القرار للمستخدم.

حققت أنظمة التوصية وحيدة المعيار (التقليدية) نجاحاً كبيراً في عدد من التطبيقات، ولكن البحث في أنظمة التوصية متعددة المعايير لم يصل بعد إلى المراحل المتطورة التي وصلتها تلك الأنظمة التقليدية، ولكي نحصل على الفائدة الكلية من التقييمات متعددة المعايير، لا بد من تطوير تقنيات توصية جديدة.

تعمل نظم التوصية التقليدية بشكل أساسي على قياس التشابه بين المستخدمين باستخدام التقييمات المعطاة من هؤلاء المستخدمين للمواد المدروسة، حيث تعتمد على نظام تقييم أحادي المعيار، أي أن المستخدم يقوم بتقييم المادة عن طريق قيمة وحيدة (تمثل التقييم الكلي)، بينما تقوم الأنظمة المتقدمة بتفصيل حاجيات المستخدم وتفضيلاته إلى عدة معايير تعبر عن تفضيل المستخدم للمادة بشكل تفصيلي بعيداً عن التقييم الكلي الذي لا يعبر بشكل جيد عن تفضيل المستخدم للمادة.

نطرح في هذه الأطروحة دراسةً مرجعية لأنظمة التوصية التقليدية وحيدة المعيار، ثم ننتقل إلى الدراسة الأساسية في الأطروحة وهي أنظمة التوصية متعددة المعايير وكيف يمكن تطبيق أنظمة دعم القرار متعدد المعايير حتى يمكن صياغة مسألة التوصية كمسألة دعم قرار للمستخدم المستهدف. يمثل الحل المقترح طريقة جديدة في حل مسائل التوصية متعددة المعايير عن طريق دراسة ارتباط المعايير ببعضها البعض وثم تشكيل نموذج يستفيد من هذا الارتباط لنستطيع بناء نموذج يكامل بين الطرق التقليدية مثل الترشيح التعاوني (المعتمد على تشابه المستخدمين) والطريقة المقترحة وهي سَلسَلة المعايير.

الكلمات المفتاحية: أنظمة التوصية الرقمية، أنظمة التوصية وحيدة المعيار، أنظمة التوصية متعددة المعايير، نظم دعم القرار متعددة المعايير.

Abstract

Recommendation systems, or also known as recommendation engines, have become an important research area and are applied in various fields. In addition, the technologies behind these systems have been developed over time. In general, these systems help users to find products or services (such as books or music) through analyzing and aggregating other users' activities, primarily in form of reviews, and then forming recommendations. Recommendations help facilitate the user's decision-making process .

While single-criterion (traditional) recommendation systems have been very successful in a number of applications, the research for multi-criteria recommendation systems is still not in its mature stages. In order to obtain the overall benefit from multi-criteria ratings, new recommendation techniques need to be developed.

Traditional Recommendation systems mainly measure the similarity between users using the reviews given by these users to the studied items. Traditional systems operate on a single-criterion rating system, i.e, the user evaluates the item by a single value (representing the overall rating). Advanced systems have detailed user needs and preferences to several criteria that express the user's preferences for the item in detail, away from the overall rating that does not reflect well the user's preferences for the item .

In this thesis, we present a state-of-the-art study of the traditional single-criterion recommendation systems. We then turn to the main study of the thesis, which is, the multi-criteria recommendation systems, and how multi-criteria decision support systems can be applied so that the recommendation can be formulated as a decision support problem for the target user. The proposed solution represents a new way to solve the multi-criteria recommendation problems by studying the correlation of criteria and then forming a model that benefits from this correlation so that we can build a model that integrates traditional methods such as collaborative filtering (based on user similarity) and the suggested method which is "Criteria Chaining."

Keywords: recommendation systems, single-criterion recommendation systems, multi-criteria recommendation systems, multi-criteria decision support systems.

الفصل الأول: مدخل لموضوع البحث وأهميته

1.2 تمهيد

إنّ النمو الكبير الذي تشهدهُ التّجارة الإلكترونيّة تركَ أثراً كبيراً على عادات المستهلكين، مزوّدةً إيّاهم بتشكيلةٍ ضخمةٍ من المنتجات، ممّا جعلها تجارة المليارات، وأيضاً مما عقّدَ الأمور عليهم لاختيار الحاجيّات الأكثر ملاءمة لأذواقِهم. كلّ ذلكَ جعلَ من وجود أنظمة التّوصية ضرورةً أساسيّةً للتعامل مع هذا الكمّ الكبير من المعلومات المتاحة، حيث تقومُ هذه الأنظمة بتقديم توصياتٍ واقتراحاتٍ للمنتجات الأكثر ملاءَمة لكلِّ شخص. دخلت هذه الأنظمة في تجارة الكتب، الأفلام، الموسيقى، المنتجات الاستهلاكيّة، الأخبار وغيرها. ويمكن تعريف مشكلة التّوصية بأنّها تقدير استجابة المستخدم للأغراض الجديدة بناءً على تاريخه القديم، وبالتالي اقتراح الأغراض التي قد تنال إعجابه منها.

وفّر التطور الأخير في علوم تكنولوجيا المعلومات العديد من موارد المعطيات الخاصة بكل شيء تقريباً. على الرغم من أن توفر المعطيات بهذا الشكل الكبير مفيد في العديد من القطاعات، إلا أنه يجعل من عملية اتخاذ القرار أكثر صعوبةً.

يملك المستخدمون والزبائن العديد من الخيارات التي قد تربكهم في عملية انتقاء الخيار المتاح الأفضل أو الأكثر مناسبةً لهم. في هذا السياق، من المهم جداً أن يتم ترشيح المعلومات وتخصيصها لاستخدام كل مستخدم على حدة.

تعتبر أنظمة التوصية من أهم طرق تخصيص اقتراحات المواد للمستخدم بناءً على حاجاته وتفضيلاته. في الأيام الحالية، تعد أنظمة التوصية الرقمية من أكثر الأنظمة المستخدمة في العديد من التطبيقات مثل تطبيقات السفر والسياحة ومشاهدة الأفلام وقراءة الكتب والاستماع إلى الموسيقى.

في تطبيقات السفر والسياحة مثلاً، يكون من الصعب والمحيّر للسائح أن يقرر ويختار وجهته من بين العديد من الوجهات المتاحة وخاصةً تلك الأماكن غير المكتشفة من قبله سابقاً [1]. كما يكون من الصعب على المشاهد أن يختار فيلماً يوافق تفضيلاته أو يختار الفيلم الأفضل لحالته، وتتكرر الحالة في التطبيقات الأخرى عندما يود المستخدم أن يقرأ كتاب أو يستمع إلى مقطوعة موسيقية أو يقوم بشراء منتج.

بالنظر للحالات السابقة والعديد من الحالات الأخرى، تتميز تقنيات استرجاع المعلومات والتقنيات المساعدة على اتخاذ القرار بشكل واسع كأدوات قيّمة في هذا المجال.

2.2 دوافع البحث

تعتمد العديد من أنظمة التوصية الرقمية على تقنيات بسيطة تقارن في العموم بين الملف الشخصي للمستخدم مع بعض الميزات الخاصة بالمواد (وجهات سياحية أو منتجات أو أفلام أو مقاطع موسيقية) [2] [1]. في مثل هذه الأنظمة، يتم سؤال المستخدم مجموعة من الأسئلة (قد تكون على شكل إعدادات الملف الشخصي) لتزويد النظام بما يقابل اهتماماته واحتياجاته وقيوده، ليقوم النظام بعملية التوصية بناءً على ربط أجوبة المستخدم مع المواد المتوفرة. تسمى هذه الأنظمة **بالأنظمة المعتمدة على المحتوى Content-Based Recommenders** [3].

تعتمد مقاربة أخرى على الحصول على معلومات مفيدة عن المستخدمين الآخرين بالتطبيق الذين يملكون صفات مشابهة أو مشتركة مع المستخدم الهدف. تستفيد الأنظمة في هذه الحالة من حقيقة أن المستخدمين المتقاربين سيكون لديهم اهتمامات وحاجات مشتركة. تسمى هذه الأنظمة **بالأنظمة المعتمدة على الترشيح التعاوني Collaborative Filtering Recommenders** [4].

يقدر المسؤول عن المنتجات في شبكة Netflix الأميركية التي تعمل عبر الإنترنت أن ثمن محرك التوصية الرقمية لديهم هو مليار دولار أميركي [5]. كما يقول الباحثون في Netflix أن الفترة التي يملكونها لشد انتباه المشاهد هي 90 ثانية على الأكثر وإلا هم معرضون لخسارة هذا الشخص [5].

3.2 إشكالية البحث

تعتمد العديد من أنظمة التوصية الرقمية على معيار وحيد يعبر عن تقييم المستخدم للمادة. إن الاعتماد على تقييمات متعددة للمادة يعطي معلومات أكثر عن تفضيل المستخدم للمادة وهو ما اعتمده موقع Trip Advisor [1] الشهير في تصميم نظام التقييم لديه. يوضح المثال التالي مشكلة الاعتماد على تقييم وحيد للفنادق.

المستخدم	المريديان	الشيراتون	الشام	أمية
محمد	3 ،2 ،5 ،2 ،3 ،2	3 ،2 ،3 ،2 ،5 ،2	3 ،2 ،5 ،3 ،2 ،2	3 ،2 ،5 ،3 ،2 ،2
علي	3 ،2 ،3 ،2 ،2 ،5	3 ،2 ،5 ،2 ،3 ،2	3 ،2 ،2 ،3 ،2 ،5	3 ،5 ،2 ،2 ،3 ،2
مجد	3 ،2 ،5 ،2 ،3 ،2	3 ،2 ،5 ،2 ،3 ،2	3 ،2 ،5 ،3 ،2 ،2	3 ،2 ،5 ،2 ،3 ،2

[1] https://www.tripadvisor.com/

تمثل الأرقام الموجودة تقييم المستخدم للمعايير التالية على الترتيب (الخدمة، المكان، النظافة، الجو العام) ويمثل الرقم الأخير الوسطي الحسابي للمعايير الأربعة.

تبرز المشكلة عند الاعتماد على تقييم وحيد (القيمة الخامسة في خلايا الجدول السابق الممثلة للتقييم الكلي) في اكتشاف النظام لتشابه المستخدم الهدف وهو محمد لكل من علي ومجد على التساوي على اعتبار أن التقييمات مشابهة بشكل كامل بين المستخدمين الاثنين لمحمد.

لدى الاعتماد على التقييمات التفصيلية تكتشف عدم موافقة علي لمحمد في آراءه التفصيلية بينما نرى أن مجد مشابه بشكل تام لمحمد.

إن الاعتماد على المعايير التفصيلية يزيد من دقة نظام التوصية الرقمي. ندعو هذه المقاربة **بأنظمة التوصية متعددة المعايير Multi Criteria Recommender Systems**.

4.2 فكرة الحل المقترح

نعتمد في حلنا في هذه الأطروحة على فكرة انتقاء معايير التقييم واستنتاج التقييم النهائي منها، حيث تقوم أنظمة التوصية متعددة المعايير بتطبيق الخوارزميات التقليدية (مثل الترشيح التعاوني) على المعايير التفصيلية واستنتاج التقييم الإجمالي لهذه التقييمات التفصيلية بناءً على تابع تجميعي Aggregation Function أو باستخدام تقنيات أخرى (مثل التقنيات الإحصائية أو تعلم الآلة).

تقوم فكرة الحل الأساسية على فكرة أن ترتيب المعايير مهم جداً (على اعتبار أن علامة التقييم التفصيلية قد تتأثر بالعلامات التفصيلية السابقة لذات المادة)، حيث نقوم باستنتاج التقييم الأول على المعيار الأول ثم نعتمد على أن هذا التقييم أصبح دخلاً في عملية توقع التقييم الثاني على المعيار الثاني وهكذا...

الفصل الثاني: الإطار النظري العام

1.3 مقدمة

جرى تعريف وتحديد مشكلة التوصية على أنها الطريقة لمساعدة الأشخاص أو المستخدمين في مجتمع معين (ما يمثله النظام الرقمي في عالم تكنولوجيا المعلومات) لإيجاد المعلومات أو المواد التي تكون غالباً مهمة أو مفيدة لهؤلاء الأشخاص أو مرتبطة باحتياجاتهم [7] [6].

يفترض نظام التوصية التقليدي وجود مجموعة من المستخدمين Users التي تمثل جميع المستخدمين في النظام، ومجموعة من المواد Items التي تمثل جميع المواد القابلة للاقتراح لهم. كما يفترض النظام وجود تابع منفعة Utility Function يقيس ملائمة اقتراح المادة i المنتمية إلى مجموعة المواد Items للمستخدم u الذي ينتمي إلى مجموعة المستخدمين Users. يعرف هذا التابع بالصيغة $R: Users \times Items \rightarrow R_0$ حيث R_0 هو عدد صحيح (أو حقيقي) موجب تماماً ينتمي إلى مجال معين [6].

من المسلَّم أن تابع المنفعة R غير معروف لكامل المجال $Users \times Items$، حيث يكون معرّف لمجال جزئي منه فقط. لذلك نريد في سياق التوصية أن نستطيع القيام بالنقاط التالية لكل مستخدم $u \ni Users$:

1- تقدير أو تقريب تابع المنفعة $R(u,i)$ لكل المواد التي يكون فيها $R(u,i)$ غير معروفاً بعد.

2- اختيار مادة أو مجموعة من المواد i التي تجعل مقدار تابع المنفعة $R(u,i)$ أعظمياً أي:

$$\forall u \in Users, i = \arg \max_{i \in Items} R(u,i) \qquad \text{Eq1}$$

يكون تابع المنفعة في معظم أنظمة التوصية الرقمية هو عبارة عن قيمة وحيدة المعيار (قيمة سلّمية)، أي التقييم الكلي لمادة ما من قبل مستخدم ما. تم اعتبار هذا الافتراض محدوداً في مراجع كثيرة [9] [8] [6]، وسبب ذلك يعود إلى أن ملاءمة المادة المقترحة لمستخدم معين قد تعتمد على أكثر من جانب متعلق بالمنفعة يأخذهم المستخدم الهدف بعين الاعتبار عند اتخاذ قراره.

تحديداً في الأنظمة التي تكون فيها التوصية معتمدة على آراء المستخدمين الآخرين (ما سميناه سابقاً بتقنية الترشيح التعاوني)، يكون دمج عدة معايير مؤثرة في رأي المستخدم مفيداً ويقود إلى نتائج أكثر دقةً.

إن المعلومات الإضافية التي تقدمها التقييمات متعددة المعايير قد تفيد في تحسين جودة الاقتراحات كونها تقدم تمثيلاً أكثر تعقيداً وتكاملاً لأفضليات المستخدم. لنأخذ المثال التالي لتوضيح الفكرة. في نظام توصيات أفلام وحيد المعيار تقليدي، يعطي المستخدم u تقييمه للفيلم i الذي شاهده مسبقاً عن طريق رقم وحيد يعبر عنه عبر $R(u,i)$. تحديداً، لنفترض أن نظام التوصية يقدم الاقتراحات عن طريق توقع تقييم المستخدم للأفلام التي لم يشاهدها بعد بناءً على تقييمات الأفلام من قبل المستخدمين الذين يملكون تفضيلات مشابهة للمستخدم الهدف، والذين نعبر عنهم عادةً بـ "**جوار المستخدم**" [10].

إن عملية تحديد جوار المستخدم هي عملية أساسية ومهمة للحصول على توقعات دقيقة للاقتراحات. على سبيل المثال، إذا شاهد مستخدمان u, u' ثلاثة أفلام، وقام كلاهما بالتعبير عن مدى رضاهما عن الأفلام الثلاثة بمقدار 6 من 10، نعتبر هذين المستخدمين جارين بين بعضهما ويتم الاعتماد على تقييمات المستخدم u' للحصول على توقع تقييمات الأفلام غير المشاهدة من قبل المستخدم u.

في المقابل، في نظام تقييم متعدد المعايير، يمكن للمستخدمين أن يعبروا عن تفضيلاتهم باستخدام تقييمات لعدة مواصفات خاصة بالمادة. في نظام توصية أفلام بمعيارين اثنين، يُسمح للمستخدم أن يقيم واصفتين للفيلم وهما (القصة والمؤثرات البصرية). يمكن لمستخدم ما أن يفضل القصة ولا يفضل المؤثرات البصرية أي $R(u,i) = (9,3)$. إذا استخدمنا تقييمين اثنين بنفس الوزن في حساب الاقتراحات، فإن التقييم الوسطي 6/10 قد يعبر عن أكثر من تركيبة من التفضيلات على مستوى المعايير الثنائية مثل $(4,8), (6,6), (9,3)\ldots$

لذلك، وعلى الرغم من أن التقييم الكلي هو 6/10 فإن المستخدمين u, u' قد يظهرا نمطين مختلفين من التقييمات التفصيلية على كل معيار. فمثلاً يعطي المستخدم u التقييمات$(9,3), (9,3), (9,3)$ للأفلام الثلاثة على التوالي، فيما يعطي المستخدم u' التقييمات $(3,9), (3,9), (3,9)$ لنفس الأفلام الثلاثة السابقة. تساعد هذه المعلومة الإضافية على تفضيلات كل مستخدم تساعد على نمذجة تفضيلات المستخدم بشكل أدق. يجب تطوير تقنيات جديدة في أنظمة التوصية الرقمية لاستغلال هذه المعلومات الإضافية بشكل جيد.

نعرض في هذا الفصل بدايةً مسألة التوصية بإطارها التقليدي وحيد المعيار، ثم ننتقل إلى عرضها كمسألة صنع قرار متعددة المعايير ونطرح إطار نظم دعم القرار متعددة المعايير في نظم التوصية الرقمية.

2.3 أنظمة التوصية الرقمية التقليدية وحيدة المعيار

تعتبر أنظمة التّوصية الرقميّة موضِعَ بحثٍ هام منذ ظهور أوّل الأوراق البحثيّة في مجال التّوصية اعتماداً على التّرشيح التّعاونيّ في أواسط تسعينيّات القرن الماضي [11]. منذ ذلك الوقت ظهرت الكثير من الأعمال لتطوير منهجيّات جديدة لأنظمة التّوصية سواءً على الصّعيد البحثيّ أو على الصّعيد التّقنيّ والتّجاريّ، حيث لا يزال الاهتمام بهذا المجال كبيراً نظراً لارتباطه بالتّطبيقات العمليّة التي تساعد المستخدمين في التعامل مع الكمّ الزّائد من المعلومات وتقديم توصيات خاصّة لكلّ منهم، كتوصية الأفلام، الأغاني، الكتب، المقالات، وغيرها...

نستعرض فيما يلي أبرز المنهجيّات المتّبعة في أنظمة التّوصية الحاليّة، مبيّنين نقاط الضّعف والقوّة في كلٍّ منها.

3.3 أنظمة التّوصية القائمة على المحتوى

إن المبدأ الأساسيّ الذي تعتمده هذه الأنظمة، هو تحديد الخواص المشتركة للأغراض التي نالت إعجاب مستخدم ما u، ومن ثمّ توصية أغراض جديدة لهذا المستخدم والتي تشترك بالخواص السّابقة مع باقي الأغراض التي أعجبت u.

يُفترض في هذه الأنظمة وجود معلومات غنيّة تصف طبيعة كل غرض i، وتتوافر هذه المعلومات بشكل شعاع من الخواص المميّزة x_i، ففي حال كانت الأغراض مستندات نصيّة كالمقالات الإخباريّة أو مستندات ويب، يكون هذا الشّعاع مؤلّفاً من أوزان Term Frequency-Inverse Document Frequency (TF-IDF) [12] لأهم الكلمات المفتاحيّة المتضمّنة في النصّ. يتم في هذه الأنظمة ربط كل مستخدم u بشعاع x_u يعبر عن الخواص المفضّلة لهذا المستخدم، حيث يمكن الحصول على هذا الشّعاع من مجموعة الأغراض التي قيّمها المستخدم I_u. تستخدم هذه الأنظمة تقنيات لإيجاد هذه الأشعّة المميّزة للمستخدمين، مثل خوارزميّة روخيو Rocchio algorithm [13]، وفيها يتمّ تحديث قيم الشّعاع x_u للمستخدم u في كلّ مرّة يقوم فيها هذا المستخدم بتقييم غرض جديد i وذلك من خلال إضافة أوزان x_i إلى x_u وفق العلاقة التاليّة:

$$x_u = \sum_{i \epsilon I_u} r_{ui}\ x_i \qquad \text{Eq2}$$

حيث r_{ui} هو تقييم المستخدم u للغرض i.

بعد الحصول على هذه الأشعّة الممثّلة لتفضيلات المستخدمين، يمكن استخدامها لتوصية أغراض جديدة تتوافق مع ميولهم من خلال اقتراح الغرض i الممثل بشعاع الخواص x_i الأكثر تشابهاً مع الشّعاع x_u الممثّل للمستخدِم، وذلك باستخدام معايير التّشابه كمعيار Cosine Similarity. [14]

تعاني أنظمة التّوصية القائمة فقط على المحتوى عادة من المشاكل التّالية: [11]

- **محدوديّة تحليل المحتوى Limited content analysis**: تعاني هذه الأنظمة من محدوديّة الخواص المرتبطة صراحة بالأغراض المطلوب توصيتها. حيث أنّه للحصول على خواص كافية لابدّ للمحتوى أن يكون إمّا قابلاً للتحليل التّلقائي باستخدام الحاسوب مثل النصوص، او أن يتمّ إدخال هذه الخواص يدويّاً للنّظام. تنفع تقنيّات استرجاع المعلومات في استخراج الخواص المميّزة للنّصوص، ولكن في بعض المجالات الأخرى يكون من الصعب استخراج هذه الخواص، كالاستخراج التّلقائي لخواص ملفّات الوسائط المتعدّدة مثل الفيديوهات والصّور الرقميّة وملفّات الصّوت وغيرها، كما أنّه من الصّعب إسناد هذه الخواص يدويّاً نتيجة لشحّ الموارد. بالإضافة إلى ما ذكر سابقاً تظهر مشكلة عدم القدرة على التمييز بين غرضين لهما مجموع الخواص نفسها، وعليه لن تتمكّن هذه الأنظمة من التّمييز بين مقالتين (جيّدة ورديئة الصّياغة) طالما أنّهما تستخدمان المصطلحات نفسها.

- **المبالغة في التّخصيص Over-specialization**: عندما تقوم هذه الأنظمة بتوصية الأغراض التي تحقّق تشابه أعلى مع الخواص التي تمّ إسنادها للمستخدم، سيكون المستخدم محدوداً بأن يتم توصيته بأغراض مشابهة للأغراض التي قام بتقييمها مُسبقاً. فعلى سبيلِ المثال، لن يتمكّن النّظام من توصية حتّى أهمّ المطاعم الإيطاليّة لمستخدم ما في حال لم يكن للمستخدم أيّ تجربة سابقة ضمن النّظام مع المأكولات الإيطاليّة. وهو ما اقترح البعض إضافة شيء من العشوائيّة للتّخلص من هذه المشكلة. وبالإضافة إلى ما ذكر سابقاً تندرج مشكلة أخرى ضمن هذا الإطار وهي أنّه في بعض الحالات ليس من المفيد توصية الأغراض ذات التّشابه الكبير مع غرض سبق للمستخدم استعراضه، مثل مقالتين تصفان الحدث ذاته والذي سبق للمستخدم القراءة عنه. وعليه اقترحت بعض أنظمة التوصية مثل DailyLearner [15]، أن تتم ترشيح الأغراض التي تشابه غرض سبق للمستخدم استعراضه في حال تجاوزت نسبة التّشابه عتبة معيّنة.
- **إقلاع المستخدم الجديد المتعثّر cold start**: يجب على المستخدم أي يقوم بتقييم عدد كافي من الأغراض قبل أن تتمكّن هكذا أنظمة من استكشاف ما يفضّله هذا المستخدم وبالتالي اقتراح أغراض جديدة تلائم رغباته. وعليه فإن المستخدم الجديد بعددٍ قليلٍ من التقييمات للأغراض ضمن النّظام لن يتمكّن من الحصول على توصيات دقيقة.

4.3 أنظمة التّوصية القائمة على مبدأ التّرشيح التّعاونيّ Collaborative Filtering

تقوم هذه الأنظمة بتوصية الأغراض اعتماداً على التّشابه بين المستخدمين أو بين الأغراض نفسها، حيث يتم توصية المستخدم بالأغراض المفضّلة للمستخدمين الأكثر تشابهاً معه. سرعان ما انتشرت هذه التّقنيّة نتيجة لنجاحها في عصر الإنترنت حيث تُستخدم هذه التّقنيّة في أنظمة Amazon.com أكبر متجر كتب عبر الإنترنت، CDNow.com أكبر متجر للأقراص المضغوطة عبر الإنترنت، وكذلك MovieFinder.com أحد أكثر مواقع الأفلام زيارة من قبل المستخدمين حول العالم. [11]

تعمل هذه الأنظمة من خلال تجميع تقييمات المستخدمين للأغراض ضمن مجال معيّن، حيث يقوم المستخدمون بتقييم الأغراض التي سبق واطّلعوا عليها لتسهيل عمليّة الانتقاء بالنسبة لغيرهم، ومنه فإنّ هذه التقييمات تحمل في جوهرها نوعاً من التّوصية الكامنة والتي يمكن الاستفادة منها في هذه الأنظمة لخلق توصيات شخصيّة مفيدة.

تكمن الفكرة الأساسيّة باحتماليّة أن يكون تقييم مستخدم ما u لغرض معيّن i مشابهاً لمستخدم آخر v، في حال كان المستخدمين السابقين قد تشابها في تقييمهما لأغراض سابقة.

تمتازُ هذه الأنظمة بإضافتها لثلاث ميزات مفتاحيّة تتلخّص في: [11]

1. القدرة على ترشيح الأغراض ذات المحتوى الذي يصعب تحليلهُ تلقائيّاً كملفات الوسائط المتعدّدة.

2. القدرة على ترشيح الأغراض بناءً على جودتها، شعبيتها ودرجة ملاءَمتها، فمثلاً تستطيع أنظمة التّوصية القائمة على المحتوى تحديد المقالات المناسبة لشخص رياضيّ، ولكنّها لا تستطيع استبعاد المقالات رديئة الصّياغة، وهو الأمر الذي يميّز أنظمة التّرشيح التّعاونيّ.
3. القدرة على تقديم توصيات مفاجئة، خارج الإطار المألوف للمستخدم (مثلاً: في حال نظام توصية أفلام، يمكن أن يكون المستخدم من الّذين يفضّلون النمط الكوميدي، ويتمكّن النّظام من توصية فلم رومانسي نتيجة تشابهه مع مستخدم آخر).

يمكن تجميع هذه الأنظمة ضمن مجموعتين (معتمد على النّماذج model-based، معتمد على الجوار neighborhood)، بالنسبة للأنظمة المعتمدة على الجوار فيتمّ استخدام تقييمات المستخدمين للأغراض المخزّنة في النّظام مباشرة لتوقّع التقييمات الجديدة، ويتحقّق ذلك بإحدى طريقتين اعتماداً على المستخدم أو اعتماداً على الغرض. أمّا بالنسبة للأنظمة المعتمدة على النّماذج ففيها يتمّ توليد نموذج خاصّ لكلّ نظام، بحيث يتم استخدام هذا النموذج في توصية أغراض جديدة للمستخدمين ملائمة لرغباتهم. يتمّ بناء النّموذج في مرحلة مبكرة قبل البدء بالتّوصيات، تُسمّى هذه المرحلة مرحلة التدريب وفيها تستخدم مجموعة من المعطيات لتكوين هذا النموذج، والّذي قد يكون شجرة قرار مثلاً.

تتنبّأ الأنظمة التي تعتمد على المستخدم بتقييم مستخدم u لغرض i بالاعتماد على تقييمات مستخدمين للغرض نفسه، يطلق على هؤلاء المستخدمين اسم الجوار ويملكون نفس تفضيلات المستخدم المراد التنبّؤ بتقييمه، ويمكن تعريف الجوار بأنّه مجموعة المستخدمين v الأكثر تشابهاً مع u وذلك من حيث تشابه التقييمات التي نسبوها للأغراض المشتركة بينهم. ومن جهة أخرى في حالة الاعتماد على الغرض، يمكن التنبّؤ بتقييم مستخدم u لغرض i اعتماداً على تقييم u لأغراض مشابهة للغرض i، بحيث يعتبر غرضان متشابهان في حال أجمع عدّة مستخدمين في النّظام على تقييمهم بالجودة نفسها.

يحوي الملحق (1) على المزيد من الدراسات المتعلقة بأنظمة التوصية الرقمية وحيدة المعيار.

ننتقل في الفقرة التالية إلى تعريف مسألة التوصية الرقمية كمسألة دعم قرار متعددة المعايير.

5.3 التوصية الرقمية كمسألة صنع قرار متعددة المعايير

يمكن أن نتّبع أحد منهجيات صنع القرار متعدد المعايير لتقديم تعدد المعايير في التوصية الرقمية. لتسهيل النقاش حول استعمال طرق وتقنيات صنع القرار متعدد المعايير عند تطوير نظام توصية رقمي، نقوم باتّباع الخطوات والأساسيات التي وضعها بيرنارد روي Bernard Roy[2] (وهو يعتبر أحد عرّابي دعم القرار متعدد المعايير في ستينيات القرن الماضي) في منهجية النمذجة العامة لمسائل دعم القرار متعدد المعايير [16]. وردت مسألة صنع القرار متعدد المعايير في الكثير من الأدبيات الأخرى التي وضحت منهجيات أخرى لنمذجة هذه المسألة [17] [18] [19]. نعتمد في بحثنا هذا على منهجية روي لتصميم نظام دعم قرار متعدد معايير.

تتضمن منهجية روي [16] أربع خطوات عند تحليل مسألة دعم قرار متعدد المعايير:

1- **تعريف غرض القرار** object of decision. أي تحديد مجموعة البدائل (المواد) التي سيتم اتخاذ القرار بناءً عليها (انتقاء أحد البدائل).
2- **تعريف عائلة متسقة من المعايير**. أي تحديد مجموعة من الوظائف التي تحدد تفضيلات صانع القرار (المستخدم الهدف) بناءً على مجموعة البدائل المختلفة. يجب أن تغطي هذه المجموعة كل المتغيّرات والمحددات التي تؤثر في قرار التوصية كما يجب أن تكون شاملة وغير زائدة عن الحاجة.
3- **تطوير نموذج تفضيل عام**. أي تعريف التابع الذي يجمّع التفضيلات الجزئية القائمة على كل معيار ضمن نموذج يحدد التفضيل الكلي لصانع القرار بخصوص بديل مرشح.
4- **اختيار إجراء دعم القرار**. وهذا يشمل تصميم وتطوير الإجرائية أو النماذج أو الأنظمة البرمجية التي ستساعد وتدعم صانع القرار عند اتخاذ قرار يخص مجموعة البدائل المتاحة، في ضوء نتائج الخطوات السابقة.

نشرح باختصار الخطوات الأربع السابقة ضمن الفقرات التالية، وكيف يمكن ربط كل خطوة منهم مع مسألة التوصية الرقمية. قبل الخوض في شرح كيفية تطبيق الخطوات السابقة ضمن نطاق التوصية الرقمية علينا أن نذكر باستخدام مصطلح "بدائل Alternatives" في سياق مسألة دعم القرار، بينما نستبدلها بمصطلح "مواد Items" في مسألة التوصية الرقمية.

[2] http://www.lamsade.dauphine.fr/~roy/

1.5.3 غرض القرار

في أنظمة التوصية الرقمية، يكون غرض القرار هو المادة i التي تنتمي إلى مجموعة المواد المتوفرة $Items$. للتعبير عن المنطق خلف القرار، يشير "روي" إلى مفهوم إشكاليات القرار Decision Problematics. يتم تحدد أربع إشكاليات أساسية:

1. **الخيار Choice**: والذي يعنى باختيار بديل أو أكثر من بديل التي يمكن اعتبارها أكثر ملاءَمة من غيرها في مجموعة البدائل.
2. **الفرز Sorting**: والذي يعنى بتصنيف البدائل إلى عدد من التصنيفات المعرفة مسبقاً.
3. **الترتيب Ranking**: والذي يشمل ترتيب كل البدائل من الأفضل حتى الأسوأ.
4. **الوصف Description**: والذي يهتم بتوصيف كل بديل في سياق مدى أدائه على كل معيار.

يمكن اعتبار جميع الإشكاليات السابقة على أنها صالحة في مسألة التوصية الرقمية:

1. اختيار وتوصية أحد أو مجموعة من المواد على أنها أكثر ملاءَمة لمستخدم معين.
2. تصنيف جميع المواد ضمن تصنيفات معرفة مسبقاً حسب ملاءمتها للمستخدم مثل "مواد مقترحة للشراء" و"مواد مقترحة للمشاهدة".
3. ترتيب جميع المواد من الأكثر ملاءمة إلى الأقل ملاءمة لمستخدم معين، وتقديم هذه اللائحة المرتبة للمستخدم.
4. شرح ملاءمة مادة معينة لمستخدم معين، بناءً على تقييمها على كل معيار على حدة. يتوافق هذا الشرح مع تحليل كامل لأداء المادة على جميع المعايير، لتوضيح ملاءمة هذه المادة للمستخدم (مما يساعد المستخدم ضمن طريقة شخصية به أن يقوم بالاختيار).

2.5.3 عائلة المعايير

يتم تحليل أداء البدائل في مجموعة المواد $Items$ بناءً على مجموعة من المعايير لكل مستخدم، وذلك لنمذجة جميع خواصهم أو واصفاتهم أو تأثيرهم أو آثارهم [16]. في أنظمة التوصية الرقمية، قد تشير المعايير إلى مجموعة ميزات المادة (في حالة أنظمة التوصية المعتمدة على المحتوى) أو إلى الأبعاد المتعددة التي يتم فيها تقييم المادة.

يمكن تمثيل أي معيار c عن طريق تابع $g_c(i)$ يعبر عن تفضيلات مستخدم ما، وذلك ليكون المستخدم قادراً على اتخاذ قرار المقارنة بين مادتين i_1, i_2. أي القدرة على تفضيل المادة i_1 على المادة i_2 (حالة $g_c(i_1) > g_c(i_2)$)، أو في حالة تكافؤ المادتين أي القدرة على استبدال أي من المادتين بالآخر ($g_c(i_1) = g_c(i_2)$). يجب التأكد من أن هذه الميزات (أو الأبعاد) تشكل عائلة **متّسقة** من المعايير، للقيام بقرارات منطقية ضمن مسألة تعدد المعايير.

يمكن القول عن مجموعة من المعايير أنها متّسقة إذا حققت الخصائص الثلاثة التالية:

1. **كاملة الترتيب Monotonic**. نقول عن مجموعة من المعايير أنها كاملة الترتيب فقط إذا تحقق الشرط: من أجل كل ثنائية بدائل i_1, i_2، وعندما يكون $g_{c1}(i_1) > g_{c1}(i_2)$ من أجل معيار معين c_1، و $g_c(i_1) = g_c(i_2)$ من أجل كل المعايير $c \neq c_1$، يمكن عندها اعتبار البديل i_1 أفضل من البديل i_2.
2. **شاملة Exhaustive**. نقول عن مجموعة من المعايير أنها شاملة فقط إذا تحقق الشرط: من أجل كل ثنائية بدائل i_1, i_2، وعندما يكون $g_{c1}(i_1) = g_{c1}(i_2)$ من أجل كل معيار c، يمكن عندها اعتبار البديلين i_1 و i_2 متكافئين.
3. **غير زائدة عن الحاجة Non-redundant**. نقول عن مجموعة من المعايير أنها غير زائدة عن الحاجة فقط إذا أدّى حذف أي معيار من هذه المجموعة إلى مخالفة الخاصتين السابقتين.

يوجد أربعة أنواع من المعايير عادةً في نظم دعم القرار متعددة المعايير [20]:

1. **قابل للقياس Measurable**: أي أن المعيار قابل للقياس كمياً على سلم تقييم كمي (مثلاً سلم من 1 إلى 5).
2. **قابل للترتيب Ordinal**: أي أن المعيار قابل لأخذ قيمة وصفية على سلم وصفي (مثلاً سلم يأخذ القيم سيء، جيد، ممتاز).
3. **احتمالية Probabilistic**: أي أن المعيار يمكن أن يستخدم توزيعات احتمالية تعبر عن عدم اليقين في التقييم.
4. **عائم Fuzzy**: أي أن المعيار يتم تقييمه عن طريق احتمال انتماء التقييم إلى أحد المجالات على سلم وصفي.

من وجهة نظر عامة، يمكن استعمال عائلة المعايير لتسهيل تمثيل تفضيلات المستخدم في أنظمة التوصية الرقمية. لذلك بمكننا اعتبار جميع أنواع المعايير الأربعة يمكن استخدامها في أنظمة التوصية الرقمية متعددة المعايير، على الرغم من استعمال أنواع معينة (القابلة للقياس والقابلة للترتيب) بشكل أكبر بكثير من الأنواع الأخرى (الاحتمالية أو العائمة).

3.5.3 نموذج التفضيل العام

إن تطوير نموذج تفضيل عام (Global Preference Model) يؤمن طريقة لتجميع التقييمات على كل معيار g_c، وذلك للتعبير عن التفضيلات بين البدائل المختلفة ضمن مجموعة المواد $Items$ بالاعتماد على إشكالية القرار المختارة. تم تطوير العديد من نماذج التفضيل العام في نظم دعم القرار متعدد المعايير. يمكن تصنيف هذه النماذج إلى تصنيفات متعددة تبعاً لصيغة النموذج المستخدم وإجرائية إنشاءه. تبعاً لـ [20] و [21]، يمكن تصنيف مقاربات نمذجة التفضيل العام تبعاً للتصنيفات التالية:

1. **نماذج تركز على القيمة Value-Focused**. يتم في هذه النماذج بناء نظام لتجميع تفضيلات المستخدم على كل معيار. في مثل هذه النماذج يتم تجميع أو تركيب التفضيلات الجزئية على كل معيار ضمن تابع للقيمة الكلية ندعوه

عادةً بتابع المنفعة. تدعى هذه المقاربة عادةً "نظرية المنفعة متعددة الواصفات (Multi Attribute Utility Theory "MAUT")".

2. **نماذج الأمثَلة متعددة الأهداف Multi-Objective Optimization**. حيث يتم التعبير عن المعايير كقيود في مسألة أمثَلة متعددة الأهداف. يكون الهدف في هذه المقاربات هو إيجاد الحل الأمثل بحسب باريتو Pareto [22] لمسألة الأمثَلة.
3. **نماذج علاقات الترتيب Outranking Relations**. حيث يتم التعبير عن الأفضليات كنظام من العلاقات التراتبية بين المواد، مما يتيح التعبير عن عدم المقارنة. في هذه المقاربة، تجري مقارنة ثنائية بين كل المواد (بعضها مع الآخر)، وتكون العلاقات من نمط: "a أفضل من b"، أو "a و b متكافآن" أو "a و b غير قابلين للمقارنة" [23] .
4. **نماذج تصنيف التفضيلات Preference Disaggregation**. حيث يتم اشتقاق نموذج التفضيلات عن طريق تحليل القرارات السابقة. تعتبر هذه المقاربات أحياناً تصنيفاً فرعياً من التصنيفات المذكورة في الأعلى، بما أنها تشتق نموذجاً آخر من صيغة معينة (نموذج يركز على القيمة أو يعتمد على علاقات الترتيب) من نموذج تفضيلات سابق أدى إلى اتخاذ قرارات في الماضي. تهدف هذه النماذج إلى إنتاج القرارات المشابهة لحالات حدثت في الماضي وتم اتخاذ قرار فيها [20].

يمكن استعمال مقاربات مختلفة من جميع التصنيفات لإنشاء نموذج تفضيل عام في نظام التوصية، وذلك اعتماداً على إشكالية القرار المعتمدة وعلى البيئة التي يعمل فيها نظام التوصية.

4.5.3 إجراء دعم القرار

في هذه الخطوة، يتم الاعتماد على قرار نهائي لمسألة دعم القرار متعددة المعايير عن طريق اختيار الطريقة المناسبة من الطرق المعرفة في الخطوات السابقة. مثل مسائل دعم القرار متعددة المعايير، تحتاج أنظمة التوصية متعددة المعايير إلى استعمال طرق مختلفة للتطبيقات والمجالات المختلفة.

على الرغم من المنهجيات المذكورة سابقاً، فإن المنظور المتبع في نظم دعم القرار متعددة المعايير هو واسع الطيف وغير مقيّد عند نمذجة نظم التوصية متعددة المعايير، وذلك لأن العديد من نظم التوصية متعددة المعايير (على الرغم من عدم اعتمادها بشكل كامل للمنهجيات السابقة الذكر) يمكن ملاءمتها ضمن نطاق مسائل دعم القرار متعددة المعايير، وذلك بسبب اعتمادها على مصادر معلومات متعددة مثل الملفات الشخصية للمستخدمين ومواصفات المواد بحد ذاتها، مما يجعلها نظم دعم قرار متعددة المعايير بحكم الواقع de facto.

من المهم أن نذكر أن نظم التوصية الموجودة لا تلتزم أحياناً بإنشاء عائلة متسقة من المعايير (لا تكون شاملة أحياناً) وهو ما يدعو إلى البحث أكثر في هذه القضية [9].

6.3 إطار نظم دعم القرار متعددة المعايير في نظم التوصية الرقمية

تصنف عادةً تقنيات التوصية الرقمية بناءً على مقاربة التوصية المستخدمة إلى عدة تصنيفات:

- الاعتماد على المحتوى Content Based
- الترشيح التعاوني Collaborative Filtering
- الاعتماد على المعرفة Knowledge Based
- الأنظمة الهجينة Hybrid Approaches

تعتمد التقنيات المعتمدة على المحتوى على تحديد الاقتراحات بناءً على مواصفات المواد التي قام المستخدم بتفضيلها في الماضي [3]، فيما تعتمد تقنيات الترشيح التعاوني على تحديد الاقتراحات بناءً على المستخدمين الآخرين الذين يملكون تفضيلات مشابهة للمستخدم. تستخدم التقنيات المعتمدة على المعرفة المعلومات المتوفرة عن المستخدمين والمواد لإيجاد متطلباتهم [24]. تتطلب تقنيات الاعتماد على المعرفة وجود قاعدة معرفة قبل البدء بالعمل وهذا ما يمثل عنق زجاجة في هذه التقنيات، لكن تفيد هذه القاعدة في تجنب الإقلاع المتعثر وتبعثر المعطيات التي تعاني منها التقنيات المعتمدة على المحتوى وعلى الترشيح التعاوني كونها تستخدم التقييمات التي قام بها المستخدم للمواد فقط [10]. تدمج الأنظمة الهجينة بين التقنيات الثلاث السابقة بعدة طرق تختلف حسب النظام وتقنية الربط.

من المهم أن نلاحظ أن العبارة "متعددة المعايير" هي عبارة عامة جداً، لذلك يمكن تصنيف أنظمة التوصية حسب طبيعة تعدد المعايير فيها إلى التصنيفات الثلاثة التالية [25]:

- البحث والتصفية في المحتوى متعدد الواصفات ونمذجة التفضيلات Multi-attribute content search, filtering, and preference modeling.
- استراتيجيات التوصية متعددة الأهداف Multi-objective recommendation strategies.
- التفضيل القائم على التقييم متعدد المعايير Multi-criteria rating-based preferences.

نوضح في الفقرات التالية نظرة عامة عن التصنيفات السابقة.

1.6.3 البحث والتصفية في المحتوى متعدد الواصفات ونمذجة التفضيلات

تسمح هذه المقاربات للمستخدم بتحديد تفضيلاته اعتماداً على مواصفات معتمدة على المحتوى عبر جميع المواد، وذلك عن طريق إجراءات البحث والتصفية (البحث عن الأفلام الكوميدية فقط مثلاً) أو عبر تحديد مواصفات المحتوى المفضلة لدى المستخدم (تحديد الممثلين المفضلين مثلاً أو تحديد تصنيفات الأفلام المفضلة كالأفلام الكوميدية مثلاً). يقوم نظام التوصية بتوصية المواد للمستخدم التي تشابه تفضيلاته الصريحة أو خيارات البحث والتصفية التي قام بالتصريح عنها.

لذلك وعلى الرغم من وجود "تعدد المعايير" بسبب وجود المواصفات المتعددة للمواد، يمكن تمثيل هذه المقاربات في النماذج الموجودة مثل الأنظمة المعتمدة على المحتوى Content-based أو حتى أنظمة استرجاع المعلومات Information Retrieval التقليدية.

2.6.3 استراتيجيات التوصية متعددة الأهداف

يكون هدف العمل على أنظمة التوصية الرقمية عادةً هو تطوير خوارزميات توصية تؤمن اقتراحات **دقيقة**، حيث يتم قياس الدقة عن طريق مقاييس مختلفة مثل الخطأ المطلق الأصغري MAE أو الجذر التربيعي لمتوسط مربعات الخطأ RMSE أو الدقة Precision أو قياس الاستجابة Recall أو العديد من المقاييس الأخرى اعتماداً على مهمة التوصية.

بَيدَ أن فهم دقة التوصية قد لا يكون دائماً متوافق دائماً مع منفعة التوصية. قدم الباحثون العديد من المقاييس الأخرى بما فيها التغطية Coverage والتنوع Diversity والحداثة Novelty والصدفة Serendipity، لقياس **أداء** (وليس دقة) نظام التوصية. كنتيجة، يمكن لأنظمة التوصية الرقمية أن تستخدم **معايير أداء متعددة** عن آخذ القرار في مجموعة المواد المقترحة للمستخدم مثل استخدام الدقة والتنوع وحداثة الاقتراحات في نظام توصية شبكة Netflix [26].

في الخلاصة، طبيعة "تعدد المعايير" في هذه المقاربات لا تظهر في محاولة تمثيل تفضيلات المستخدم بشكل أكثر تعقيداً وإنما عن طريق أمثَلة عدة أهداف لأداء نظام التوصية.

3.6.3 التفضيل القائم على التقييم متعدد المعايير

تستفيد هذه التقنيات من التقييمات متعددة المعايير والتي تمثل تفضيلات المستخدم الشخصية لمكونات (أو مواصفات) مختلفة للمواد. كمثال، يمكن للمستخدم في مثل هذه الأنظمة ألا يقيم الفيلم من منطلق الرضى الكلي عنه، وإنما يمكنه أن يقيم عدة عوامل في الفيلم مثل القصة والمؤثرات البصرية والتمثيل. يمكن أن نقول عن هذه الأنظمة انها تسمح للمستخدم أن يقوم بتحديد تفضيلاته بشكل أدق عن طريق تحديد تقييمه على عدة معايير (تقييم الفيلم Avatar بـ 3\5 للقصة و5\5 للمؤثرات البصرية)، ومن ثم الاستفادة من هذه معلومات التفضيلات المصقولة في توصية المواد.

تختلف هذه المقاربة عن مقاربة المحتوى متعدد الواصفات في عدم تحديد المستخدم لتفضيلاته أو أهمية المؤثرات البصرية للأفلام واستخدامها في استعلام من قبله، وإنما يحدد المستخدم كم أحبَّ المؤثرات البصرية في فيلم محدد.

7.3 الخلاصة

في الخلاصة، وكما رأينا سابقاً، يمكن رؤية العديد من أنظمة التوصية الرقمية (التي توظّف التقنيات التقليدية مثل الاعتماد على المحتوى أو الترشيح التعاوني أو الاعتماد على المعرفة أو النظم الهجينة) على أنها نظم توصية متعددة المعايير بطريقة أو بأخرى. تقوم بعض هذه المقاربات بنمذجة تفضيلات المستخدم بناءً على المواد متعددة الواصفات التي قام المستخدم بتفضيلها في الماضي، بينما تقوم أنظمة أخرى بالاعتماد على السماح للمستخدم بتحديد تفضيلاته للمواد على هيئة بحث وتصفية ضمن هذه المواد، وتقوم باقي الأنظمة بتأمين التوصيات بموازنة العديد من مقاييس الأداء.

على الرغم من ذلك، فإن الاتجاه الحالي في نظم التوصية متعددة المعايير هو دراسة مقاربات مبتكرة في الترشيح التعاوني عبر محاولة نمذجة تفضيلات المستخدم بطريقة دقيقة وشاملة عن طريق التقييم على معايير متعددة [25].

نعتقد أن المعلومات الإضافية في تفضيلات المستخدم تقدّم فرص عديدة لتقديم نظم توصية رقمية هادفة عن طريق بناء بيئة تقييم متعددة المعايير. لذلك نقدم في الفصل القادم دراسة مرجعية للتقنيات المستخدمة في هذا النمط من أنظمة التوصية التي تستخدم تقييمات فردية على معايير متعددة، والتي سنشير إليها نظم التوصية متعددة المعايير Multi-Criteria Rating Recommenders.

الفصل الثالث: الدراسة المرجعية

1.4 نظم التوصية متعددة المعايير Multi-Criteria Rating Recommendation

نعرف في هذا الفصل مسألة التوصية ذات التقييم متعدد المعايير انطلاقاً من نظيرتها ذات التقييم وحيد المعيار، كما نناقش الفوائد التي تقدمها فكرة تعدد المعايير لأنظمة التوصية الرقمية. نعرف أولاً بالتوصية وحيدة المعيار التقليدية وثم ننطلق إلى توسعتها لتشمل التقييم متعدد المعايير.

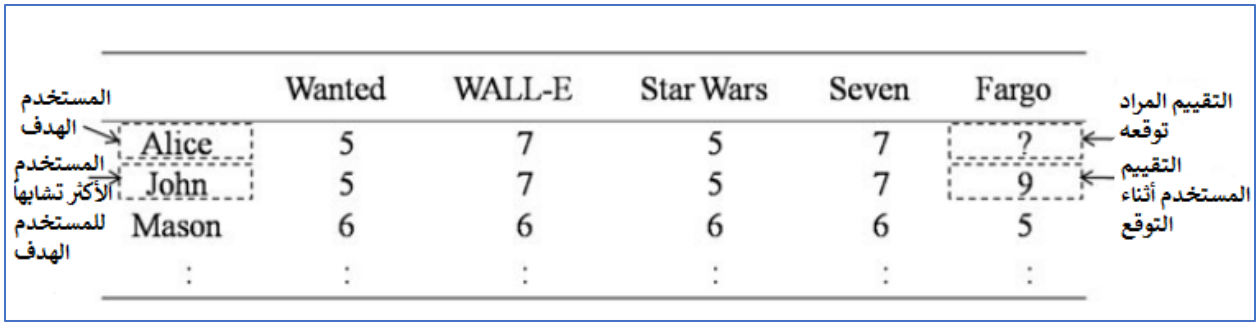
1.1.4 التوصية وحيدة المعيار التقليدية

تعمل أنظمة التوصية الرقمية التقليدية في فضاء ثنائي البعد مكون من المستخدمين والمواد $Users \times Items$. تمثل منفعة Utility المواد للمستخدمين عادةً عن طريق مجموعة مرتبة ترتيباً كلياً من التقييمات. يمكن أن تكون هذه المجموعة أحادية العنصر (مثل شراء منتج) أو ثنائية العناصر (مثل الإعجاب Like وعدم الإعجاب Dislike، أو مرتفع High ومنخفض Low، أو جيد Good وسيء Bad)، أو مجموعة محددة من القيم المرتبة (مثل سلم التقييم من نجمة واحدة إلى خمس نجوم)، أو مجال من الأرقام (مثل $[-10, +10]$) [6].

يكون التابع R معروفاً بشكل صريح لمجموعة جزئية من الفضاء $Users \times Items$، وهو مجموعة المواد التي قام المستخدمون بتقييمها سابقاً. ويكون معظم المجال $Users \times Items$ غير معروفاً. تهدف أنظمة التوصية إلى توقع منفعة (أو يمكننا استخدام "تقييم") مادة ما لمستخدم ما. يمكن كتابة التابع R بالشكل:

$$R: Users \times Items \rightarrow R_0 \qquad \text{Eq3}$$

يتم تحديد التابع R اعتماداً على مدخلات المستخدم مثل التقييمات العددية التي قام بتزويدها صراحةً للمواد و/أو بيانات المعاملات التي تظهر تفضيلات المستخدم (مثل سجل المشتريات التي قام المستخدم بشرائها). تستخدم معظم نظم التوصية الرقمية تقييمات وحيدة المعيار التي تعبر عن تفضيل المستخدم لمادة ما بشكل عام. على سبيل المثال، في نظام توصية أفلام وكما هو مبين في *الشكل 1*، قام المستخدم Alice بتقييم الفيلم Wanted بقيمة 5\10، يمكن أن نكتب $R(Alice, Wanted) = 5$. لتوضيح الفكرة نفترض أننا نستخدم تقنية الترشيح التعاوني لإيجاد جوار المستخدم ومنه توقع التقييم [4]. تتوقع هذه التقنية تقييم المستخدم لمادة معينة بناءً على تقييمات المستخدمين الآخرين الذين يملكون تفضيلات مشابهة (أي جوار المستخدم). بشكل محدد، في المثال المرفق في *الشكل 1*، يحاول نظام التوصية توقع تقييم المستخدم Alice للفيلم Fargo بناءً على التقييمات الملاحظة. بما أن المستخدمين Alice وJohn يظهران تشابه في نمط تقييماتهما على الأفلام الأربعة الأخرى التي شاهدوها سابقاً وقاموا بتقييمهما. لغرض هذا المثال البسيط، يتوقع النظام تقييم فيلم Fargo للمستخدم Alice بناءً على تقييم المستخدم John له (John هو جوار Alice في هذه الحالة) أي 9\10. يجب أن نلاحظ أننا نستخدم في الحالات العامة تقييمات من عدة مستخدمين وليس مستخدم واحد فقط في الأنظمة الواقعية.



الشكل 1 نظام توصية أفلام وحيد المعيار

2.1.4 توسعة نظم التوصية التقليدية لتشمل التقييم متعدد المعايير

مع نمو عدد من تطبيقات العالم الواقعي، جرى اعتبار توسعة تقنيات التوصية لتأخذ بعين الاعتبار التقييم متعدد المعايير على أنها واحدة من القضايا المهمة في الجيل الجديد من أنظمة التوصية الرقمية [6]. بعض الأمثلة لأنظمة التوصية ذات التقييم متعدد المعايير تشمل الأنظمة التالية:

- دليل Zagat[3] الذي يؤمن ثلاثة معايير لتقييم المطاعم (الطعام والديكور والخدمة).
- موقع Buy.com الذي يؤمن عدة معايير لتقييم الأجهزة الإلكترونية (حجم العرض والأداء ومدة حياة البطارية والكلفة).
- موقع Yahoo! Movies الذي يسمح بتقييم الأفلام على أربعة معايير (القصة والحركات والإخراج والمؤثرات البصرية).

يمكن للمعلومات الإضافية عن تفضيلات المستخدم التي تقدمها التقييمات متعددة المعايير (بدلاً من المعيار الوحيد) أن تشكل عوناً في تحسين أداء نظام التوصية الرقمي.

تختار بعض أنظمة التقييم متعدد المعايير أن تُنمذج منفعة (أو تقييم) المستخدم لمادة معينة عن طريق تقييم عام R_0 بالإضافة إلى التقييمات المختلفة $R_1, R_2, \ldots, R_k$ على كل معيار على حدة $(c = 1, \ldots, k)$ c، بينما تختار أنظمة أخرى أن تنمذج التقييم بدون استعمال التقييم الكلي R_0 وتركز فقط على التقييمات الإفرادية. لذلك يمكن تمثيل تابع المنفعة في مسألة التوصية متعددة المعايير المعتمدة إما مع التقييم الكلي أو بدونه كالتالي:

$$R: Users \times Items \rightarrow R_0 \times R_1 \times \ldots \times R_k \qquad \text{Eq4}$$

أو:

[3] https://www.zagat.com/

$$R: Users \times Items \rightarrow R_1 \times \ldots \times R_k \qquad \text{Eq5}$$

بوجود التقييمات متعددة المعايير (ومع وجود التقييم الكلي R_0)، توضح الأشكال الشكل *1* والشكل 2 الفوائد الممكنة من المعلومات الإضافية الممثلة في التقييمات الإضافية على المعايير المختلفة.

		Wanted	WALL-E	Star Wars	Seven	Fargo	
المستخدم الهدف	Alice	5,2,2,8,8	7,5,5,9,9	5,2,2,8,8	7,5,5,9,9	?,?,?,?,?	التقييمات التي يجب توقعها
	John	5,8,8,2,2	7,9,9,5,5	5,8,8,2,2	7,9,9,5,5	9,8,8,10,10	
المستخدم الأقرب للمستخدم الهدف	Mason	6,3,3,9,9	6,4,4,8,8	6,3,3,9,9	6,4,4,8,8	5,2,2,8,8	التقييمات المستعملة أثناء التوقع
	⋮	⋮	⋮	⋮	⋮	⋮	

الشكل 2 نظام توصية أفلام متعدد المعايير (التقييمات لكل فيلم: التقييم الكلي، القصة، الحركات، الإخراج، المؤثرات البصرية)

بينما يمتلك كل من Alice وJohn تفضيلات متشابهة في نظام توصية وحيد المعيار (الشكل *1*)، يمكن أن نرى في نظام توصية متعدد المعايير أنهما يظهران تفضيلات مختلفة لمختلف تفاصيل (معايير) الأفلام، على الرغم من امتلاكهما تفضيلات متشابهة في التقييمات الكلية (الشكل 2). بعد التدقيق في كل المعلومات التي تظهرها التقييمات المتعددة، يمكن أن نشاهد أن Alice وMason يظهران نمط تقييم متشابه (أكثر بكثير من النمط الذي يظهره كل من Alice وJohn). لذلك، وباستخدام نفس مقاربة الترشيح التعاوني المستخدمة سابقاً، ولكن مع أخذ التقييمات متعددة المعايير بعين الاعتبار، يمكن أن نتوقع أن يكون تقييم Alice للفيلم Fargo هو 5/10 بناءً على تقييم Mason لنفس الفيلم.

يحوي هذا المثال على أن التقييم الكلي فقط يمكن أن يخفي التفاوت الضمني في تفضيلات المستخدم للجوانب المختلفة لمادة معينة، ويمكن للتقييمات متعددة المعايير أن تساعد على فهم تفضيلات كل مستخدم، ومنه القدرة على تزويد المستخدم بتوصيات أدق. يوضح المثال أيضاً كيف يمكن للتقييمات متعددة المعايير أن تنتج توصيات موجهة بدقة، مثلاً، عن طريق اقتراح الأفلام التي تحقق أفضل تشابه على معيار القصة إذا كان هذا المعيار هو المعيار الأهم للمستخدم.

لذلك، نحتاج لتقنيات وخوارزميات توصية جديدة تستفيد من التقييمات متعددة المعايير في أنظمة التوصية. بما أن أنظمة التوصية تحسب وتقدم توصيات باستخدام إجراء ذو مرحلتين، وهما مرحلة توقع التقييم ومرحلة توليد الاقتراحات. يمكن استخدام معلومات التقييمات متعددة المعايير في كلا المرحلتين بطرق مختلفة.

تم تطوير بعض المقاربات لمرحلتي التوقع والتوصية ويوجد بعض الأنظمة التي تنفذ هذه المقاربات والتي سنقوم بتحليلها في الفقرتين التاليتين.

التوقع Prediction: وهي المرحلة التي يتم فيها حساب تفضيلات المستخدم (أي تقييماته المتوقعة). عادةً، في هذه المرحلة، يتوقع نظام التوصية تابع المنفعة R لكل أو جزء من الفضاء $Users \times Items$ اعتماداً على التقييمات المعروفة سابقاً وبعض المعلومات الأخرى (مثل الملفات الشخصية للمستخدمين و/أو محتوى المواد)، أي في كلمات أخرى، يتم حساب توقع التقييم للمواد غير المعروف.

التوصية Recommendation: في هذه المرحلة يتم استخدام التوقعات المحسوبة لدعم قرار المستخدم عن طريق إجراء توصية معين، أي هي المرحلة التي يتم فيها توصية مجموعة من N أفضل مادة التي تعطي أعظم منفعة (مثل توصية أفضل N مادة الذين يملكون أفضل تقييمات متوقعة والتي تحقق بعض الشروط الإضافية مثل تنوع وحداثة المواد).

نقوم أولاً بتصنيف التقنيات الموجودة في نظم التوصية ذات التقييمات متعددة المعايير إلى تقنيات تستخدم أثناء توقع التقييم وتقنيات تستخدم أثناء توليد التوصيات، كما نشرح هذه التقنيات بالتفصيل في الفقرتين القادمتين. يقدم الجدول 1 نظرة عامة عن هذه التقنيات.

الجدول 1 تقنيات نظم التوصية ذات التقييمات متعددة المعايير

المرحلة	التقنيات المستعملة	
توقع التقييم Rating Prediction	مقاربات تعتمد على التجربة Heuristic Based	مقاربات تعتمد على النماذج Model Based
	– استخدام التقييمات متعددة المعايير لتحسين حساب تشابهات مستخدم–مستخدم أو مادة–مادة في الترشيح التعاوني المعتمد على الجوار: o حساب قيم التشابه على كل معيار، ثم تجميع التشابهات في قيمة تشابه وحيدة (يمكن أن نستخدم أوزان أهمية لكل معيار). o حساب قيم التشابه باستخدام مقاييس مسافة متعددة الأبعاد مباشرةً على أشعة التقييمات متعددة المعايير. – توقع التقييمات بشكل تجريبي باستخدام نمذجة المنطق العائم Fuzzy Modeling: o النمذجة العائمة اللغوية Fuzzy Linguistic Modeling. o تجميع التفضيلات متعددة المعايير العائم Fuzzy multi-criteria preference modeling.	– بناء نماذج توقع لتقدير التقييمات غير المعروفة بناءً على التقييمات متعددة المعايير: o المقاربة التقليدية: بناء نماذج لتجميع معطيات التقييم الإفرادية إلى تقييم كلي واحد. – المقاربات النموذجية Representative model-base approaches: o توابع التجميع البسيطة: الوسطي البسيط، الانحدار الخطي Linear regression. o النمذجة القائمة على الاحتمالات: نماذج الخليط المرنة Flexible Mixture Models، التحليل الدلالي الاحتمالي الكامن probabilistic latent semantic analysis. – Multi-linear singular value decomposition (MSVD). – توابع التجميع المعقدة: Support Vector Regression (SVR).
توليد التوصيات (إيجاد أفضل المواد)	– عند وجود التقييم الكلي R_0 (ضمن التقييمات المتعددة): o ترتيب المواد حسب تقييمها المتوقع. – عند عدم وجود التقييم الكلي R_0: o تصميم ترتيب كلي للتوصيات عن طريق استخدام مقاربة UTA (UTilités Additives) [27]. o إيجاد الحل الأمثل من نمط Pareto. o استخدام المعايير الإفرادية كمرشحات للتوصية.	

2.4 تضمين التقييم متعدد المعايير أثناء التوقع Engaging Multi-Criteria Ratings During Prediction

نغطي في هذه الفقرة نظرة عامة عن التقنيات التي تستعمل التقييمات متعددة المعايير في توقع التقييم الكلي أو التقييمات الإفرادية (أو كليهما). في العموم، يمكن تصنيف تقنيات التوصية عبر تشكيل تابع المنفعة إلى تصنيفين: تقنيات تعتمد على التجربة Heuristic Based (يتم الإشارة إليها أحياناً بالمعتمدة على الذاكرة memory-based)، وتقنيات تعتمد على النماذج Model Based [6] [28].

تقوم التقنيات المعتمدة على التجربة بحساب منفعة (تقييم) كل مادة للمستخدم بسرعة وبشكل آلي اعتماداً على المعلومات الملاحَظة للمستخدم ومبنية بشكل أساسي على فرضية تجريبية معينة. على سبيل المثال، التقنية المعتمدة على الجوار، واحدة من أشهر تقنيات الترشيح التعاوني التي تعتمد على التجريبيات. تفترض هذه التقنية أن أي مستخدمَين يظهران تفضيلات متشابهة على المواد المشاهَدة سيظهران تفضيلات متشابهة على المواد غير المشاهَدة (غير المعروفة) كذلك. على النقيض من ذلك، تقوم التقنيات القائمة على النماذج بتعلم نموذج توقعي Predictive Model، الذي يشرح المعطيات المشاهَدة بالشكل الأفضل، ومن ثم نستعمل هذا النموذج المتعَلَّم لتوقع تقييم المواد غير المشاهَدة للتوصية.

تبعاً لهذا التصنيف، نقدم خوارزميات التوصية المعتمدة على التقييمات متعددة المعايير عن طريق تجميعهم إلى مقاربات تعتمد على التجربة ومقاربات تعتمد على النماذج.

1.2.4 المقاربات التجريبية Heuristic Approaches

جرى العمل في بعض المقاربات السابقة على توسيع حساب التشابه في تقنيات الترشيح التعاوني التقليدية المعتمدة على التجربة لتعكس المعلومات الناتجة عن التقييمات متعددة المعايير [8] [29] [30]. في هذه المقاربة، يتم حساب التشابه بين المستخدمين عن طريق تجميع التشابهات التقليدية للمعايير الإفرادية أو عن طريق مقاييس المسافة متعددة الأبعاد. يجب أن نلاحظ أن هذه المقاربة تغير في المجتزأ الخاص بحساب التشابه في خوارزميات التوصية التقليدية، أي عند حساب التشابه، تبقى إجرائية حساب التقييم الكلي نفسها.

تتوقع تقنية التوصية القائمة على الترشيح التعاوني المعتمد على الجوار التقييمات غير المعروفة لمستخدم معين اعتماداً على التقييمات المعروفة للمستخدمين الآخرين اللذين يملكون تفضيلات متشابهة لهذا المستخدم (جوار المستخدم). لذلك، فإن الخطوة الأولى من إجراءات التوقع هي اختيار طريقة حساب التشابه لإيجاد عائلة الجوار لكل مستخدم. تم استخدام العديد من الطرق لحساب التشابه في أنظمة التوصية وحيدة المعيار، ومن أشهر هذه الطرق هي الاعتماد على الترابط Correlation والاعتماد على الزوايا Cosine.

يمثل $R(u,i)$ التقييم الذي أعطاه المستخدم u للمادة i، ويمثل $\overline{R(u)}$ وسطي التقييمات للمستخدم u. على افتراض أن $I\ (u,u')$ تمثل مجموعة المواد المشتركة التي قام المستخدمان u,u' بتقييمهم، يمكن كتابة مقياسين شهيرين بالشكل:

- مقياس Pearson المعتمد على الترابط:

$$sim(u,u') = \frac{\sum_{i\in I(u,u')}\left(R(u,i) - \overline{R(u)}\right)\left(R(u',i) - \overline{R(u')}\right)}{\sqrt{\sum_{i\in I(u,u')}\left(R(u,i) - \overline{R(u)}\right)^2}\sqrt{\sum_{i\in I(u,u')}\left(R(u',i) - \overline{R(u')}\right)^2}} \qquad \text{Eq6}$$

- مقياس التشابه Cosine-based:

$$\cos(x_u, x_v) = \frac{x_u\, x_v}{\|x_u\|\,\|x_v\|} \qquad \text{Eq7}$$

$$CV(u,v) = \cos(x_u, x_v) = \frac{\sum_{i\in \mathrm{I}_{uv}} R(u,i)\, R(v,i)}{\sqrt{\sum_{i\in \mathrm{I}_u} R(u,i)^2 \ \sum_{j\in \mathrm{I}_v} R(v,i)^2}} \qquad \text{Eq8}$$

لا يمكن لأنظمة التوصية ذات التقييم متعدد المعايير أن توظف العلاقات السابقة بشكل مباشر، وذلك لأن $R(u,i)$ تحوي التقييم الكلي R_0 و k تقييم إفرادي على عدة معايير $R_1, \dots, R_k$، أي أن $R(u,i) = (R_0, R_1, \dots, R_k)$، لذلك يكون لدينا k+1 تقييم لكل زوج (u,i) بدلاً من تقييم وحيد. تم استخدام مقاربتين مختلفتين لحساب التشابه باستخدام $k+1$ قيمة تقييم في حساب التشابه بين المستخدمين.

المقاربة الأولى: تجميع التشابهات الإفرادية

تقوم المقاربة الأولى **بتجميع** التشابهات التقليدية القائمة على كل تقييم فردي. تحسب هذه المقاربة أولاً التشابه بين مستخدمَين بشكل منفصل على كل معيار منفصل، وذلك باستخدام أي تقنية لحساب التشابه التقليدي مثل التقنيات المعتمدة على الترابط أو على الزوايا. ثم نحصل على التشابه النهائي بين المستخدمَين عن طريق تجميع $k+1$ قيمة تشابه إفرادية. يطرح Adomavicius وKwon [8] مقاربتين للتجميع وهما وسطي التشابهات أو الحالة الأسوأ (التشابه الأقل) كما هو مبين في المعادلات 9، 10، 11. كمقاربة عامة، يطرح Tang وMcCalla [30] في نظامهما لتوصية الأوراق البحثية حساب التشابه على أساس مجموع مثقّل من التشابهات الفردية على معايير مختلفة للورقة البحثية (التقييم الكلي، القيمة المضافة، درجة تقييمها من القرين Peer Review، والميزات المطروحة للمتعلم مثل المعرفة السابقة والاهتمام بها). في هذه المقاربة، نختار الوزن لكل معيار c، ونعبر عنه بـ w_c، بحيث يعكس أهمية وفائدة هذا المعيار ضمن عملية التوصية.

- التشابه الوسطي:

$$sim_{avg}(u,v) = \frac{1}{k+1}\sum_{c=0}^{k} sim_c(u,v) \qquad \text{Eq9}$$

- التشابه ذو الحالة الأسوأ:

$$sim_{avg}(u,v) = \min_{c=0,\dots,k} sim_c(u,v) \qquad \text{Eq10}$$

- التشابه عبر الوسطي المثقل:

$$sim_{aggregate}(u,v) = \sum_{c=0}^{k} w_c\, sim_c(u,v) \qquad \text{Eq11}$$

المقاربة الثانية: استخدام قياسات مسافة متعددة الأبعاد

تقوم المقاربة الثانية بحساب التشابه باستخدام مقاييس مسافة متعددة الأبعاد، مثل مقاييس مانهاتن Manhattan أو الإقليدية Euclidean، أو تشيبيشيف Chebyshev [8]. يمكن حساب المسافة بين مستخدمين u,v على المادة i أي $d(R(u,i),R(v,i))$ باستخدام أحد المعادلات التالية:

- مسافة مانهاتن:

$$\sum_{\boldsymbol{c=0}}^{\boldsymbol{k}} |r_c(u,i) - r_c(v,i)| \qquad \text{Eq12}$$

- المسافة الإقليدية:

$$\sqrt{\sum_{\boldsymbol{c=0}}^{\boldsymbol{k}} |r_c(u,i) - r_c(v,i)|^{\boldsymbol{2}}} \qquad \text{Eq13}$$

- مسافة تشيبيشيف (أو المسافة العظمى):

$$\max_{\boldsymbol{c=0,\dots,k}} |r_c(u,i) - r_c(v,i)| \qquad \text{Eq14}$$

يمكن حساب المسافة الكلية بين مستخدمَين ببساطة عن طريق وسطي المسافات لكل المواد التي قيماها المستخدمان بشكل مشترك، ويمكن كتابتها على الشكل:

$$dist(u,v) = \frac{1}{|I(u,v)|}\sum_{i\in I(u,v)} d(R(u,i),R(v,i)) \qquad \text{Eq15}$$

كلما كان المستخدمان أكثر تشابهاً (أي كلما كان التشابه أكبر بينهما)، كلما كانت المسافة أقل بينهما. لذلك نحتاج التحويل البسيط التالي بسبب العلاقة العكسية بين المقياسين (المسافة والتشابه):

$$sim(u,v) = \frac{1}{1 + dist(u,v)} \qquad \text{Eq16}$$

طرح Manouselis و Costopoulou [29] ثلاث خوارزميات مختلفة لحساب التشابهات بين المستخدمين في إعدادات التقييمات متعددة المعايير: التشابه لكل أولوية والتشابه لكل تقييم والتشابه لكل منفعة جزئية.

تحسب خوارزمية التشابه لكل أولوية التشابه بين المستخدمين بناءً على أوزان $w_c(u)$ للمستخدم u لكل معيار c. في هذه الطريقة، تُنشئ الخوارزمية جوار المستخدمين الذين يملكون أولويات (أوزان) متشابهة للمعايير المختلفة مع المستخدم الهدف. ثم تقوم بتوقع الفائدة الكلية لمادة معينة بناءً على تقييمات المستخدمين في الجوار.

تنشئ الخوارزميتان الآخرتان جواراً منفصلاً للمستخدم الهدف لكل معيار، أي تقوم بحساب التشابه مع المستخدمين الآخرين على كل معيار، ومن ثم تتوقع التقييم التي يمكن أن يعطيه المستخدم على كل معيار. تحسب خوارزمية التشابه لكل تقييم التشابه بناءً على التقييم غير المثقّل التي يعطيه المستخدم لكل معيار، بينما تحسب خوارزمية التشابه لكل منفعة جزئية التشابه بناءً على التقييم المثقّل (باستخدام $w_c(u)$ لكل مستخدم u) التي أعطاها المستخدم لكل معيار.

في مثل هذه الأنظمة، نحصل على التشابهات بين المستخدمين باستخدام التقييمات متعددة المعايير، وتبقى إجرائية التوصية كما هي في أنظمة التوصية وحيدة المعيار التقليدية.

تكون الخطوة التالية من أجل مستخدم معين أن نقوم بإيجاد مجموعة الجوار الذين يملكون قيم التشابه الأعلى وتوقع التقييمات الكلية غير المعروفة للمستخدم بناءً على تقييمات الجوار. لذلك، تكون هذه المقاربات المعتمدة على التشابه قابلة للتطبيق في تقنيات التوصية المعتمدة على الترشيح التعاوني بناءً على الجوار التي تحتاج لحساب التشابه بين المستخدمين (أو المواد).

في الخلاصة، يمكن استخدام التقييمات متعددة المعايير لحساب التشابه بين مستخدمَين باستخدام أحد الطريقتين [8]: (1) تجميع قيم التشابه التي يتم حسابها بشكل منفصل على كل معيار ضمن قيمة تشابه واحدة و (2) حساب المسافة بين التقييمات متعددة المعايير بشكل مباشر ضمن الفضاء متعدد الأبعاد.

أظهرت النتائج التجريبية باستخدام عينة صغيرة من Yahoo! Movies أن كلا الطريقتين المعتمدتين في التقييمات متعددة المعايير تعطي نتائج أفضل من قرينتها في النظم وحيدة المعيار التقليدية بمقدار 3.8% باستخدام مقياس الدقة precision، الذي يمثل نسبة التقييمات الكلية الحقيقية على تلك التي توقعها النظام لتكون أعلى N مادة قريبة من كل مستخدم [8]. يعتمد التحسين في الدقة على عدة معاملات تدخل في تقنيات الترشيح التعاوني مثل أحجام الجوار وقيمة N في أعلى N مادة مقترحة للمستخدم. علاوة على ذلك، يمكن توسعة هذه المقاربات كما اقترح Manouselis و Costopoulou [29] وذلك عن طريق حساب التشابهات بتوسعة استخدام التقييمات المعروفة إلى استخدام أوزان الأهمية لكل معيار. تم تقييم المقاربات الأخيرة في

تطبيق عبر الإنترنت يقوم بتوصية الأسواق الإلكترونية للمستخدمين، حيث يدخل الباعة والمشترين إلى هذا التطبيق لتبادل المعلومات عن الأسعار والعروض، بناءً على التقييمات متعددة المعايير على الأسواق الإلكترونية. أعطت خوارزمية التشابه عبر الأولوية باستخدام المسافة الإقليدية نتائج أفضل عبر مقارباتهم المقترحة من ناحية الخطأ الوسطي بالقيمة المطلقة MAE (0.235 على سلم 1–7) مع تغطية عالية نسبياً (93% من المواد التي يمكن أن يتم اقتراحها)، بالمقارنة مع الخوارزميات غير المخصصة، مثل الاعتماد على الوسطي الحسابي التي أعطت نتائج بخطأ MAE أعلى (0.718) مع تغطية بقيمة 100% [29].

درس Maneeroj et al [31] بشكل معمق مسألة إيجاد الجوارات الأكثر مناسبةً في إعدادات التوصية متعددة المعايير. بالتحديد، اعتماداً على ملاحظة أن المعايير المختلفة قد يكون لها أوزان مختلفة للمستخدمين المختلفين، اقترحوا مقاربةً تدمج مستويات الأهمية الإفرادية لكل معيار ضمن إجراء حساب التشابه بين مستخدم وآخر. يمكن لهذه المقاربة أن تعطي جوارات مختارة بشكل أفضل وبالتالي تعطي نتائج توصية أفضل.

كما ذكرنا سابقاً، عند حساب التشابه بين المستخدمين أو المواد، تقوم خوارزمية الترشيح التعاوني المعتمدة على الجوار التقليدية بتقدير التقييم الذي يمكن أن يعطيه المستخدم u أن يعطيه للمادة i عن طريق حساب الوسطي المثقّل لكل التقييمات المعروفة $R(v,i)$، حيث v مشابه للمستخدم u. يمكن سرد طريقتين شهيرتين لحساب الوسطي المثقّل كالتالي:

- مقاربة الجمع المثقّل Weighted Sum:

$$R(u,i) = \frac{\sum_{\acute{u}\in N(u,i)} sim(u,\acute{u})R(\acute{u},i)}{\sum_{\acute{u}\in N(u,i)}|sim(u,\acute{u}|} \qquad \text{Eq17}$$

- مقاربة الجمع المثقّل المضبوط Adjusted Weighted Sum:

$$R(u,i) = \overline{R(u)} + \frac{\sum_{\acute{u}\in N(u,i)} sim(u,\acute{u})(R(\acute{u},i) - \overline{R(\acute{u})})}{\sum_{\acute{u}\in N(u,i)}|sim(u,\acute{u}|} \qquad \text{Eq18}$$

إن قيمة التقييم $R(v,i)$ تكون مثقّلة عبر التشابه بين المستخدم v مع المستخدم u، حيث يكون كلما كان المستخدمان متشابهان أكثر، كلما كان وزن $R(v,i)$ داخلاً في حساب التقييم $R(u,i)$. تمثل $N(u,i)$ مجموعة المستخدمين المشابهين للمستخدم u عبر هؤلاء الذين استخدموا المادة i، ويمكن أن يتراوح حجم المجموعة $N(u,i)$ من 1 إلى جميع المستخدمين في كل مجموعة الدراسة. تحديد حجم الجوار بعدد محدد (مثلاً 3) سيحدد عدد المستخدمين المشابهين الذي سيتم استخدامهم في عملية حساب التقييم $R(u,i)$.

في المقاربة المعتمدة على التشابه المشروحة أعلاه تمثل مقاربة معتمدة على المستخدمين تستخدم جوار المستخدمين لحساب التوصيات. كما يمكن تحويل المقاربة المعتمدة على المستخدمين باستخدام معيار وحيد بشكل مباشر إلى مقاربة تعتمد على المواد تستخدم جوار المواد لحساب التوصيات [32]، يمكن إعادة كتابة المعادلات المستخدمة في المقاربة المعتمدة على المستخدمين إلى مقاربة تعتمد على المواد بشكل مباشر.

علاوةً على ذلك، إن وجوب تحديد تقييمات رقمية دقيقة للمعايير المختلفة لكل مادة قد يشكل عبء متزايد على المستخدم. لذلك، يمكن أن يكون من المفيد أن نأخذ بعين الاعتبار الطبيعة الشخصية وغير الدقيقة والغامضة للتقييمات الإنسانية عند تجميع المعلومات. طرحت العديد من الدراسات استخدام مقاربات المعتمدة على اللغة بشكل عائم Fuzzy Linguistics لتمثيل وتجميع تقييمات المستخدم وتوظيف تقنيات صنع القرار متعددة المعايير العائمة في ترتيب وتصنيف المواد لكل مستخدم [33] [34]. بالتحديد أكثر، يمكن أن يتم تجميع آراء المستخدم ضمن شكل كيفي (أي ضمن تعابير لغوية). على سبيل المثال، في عمل Boulkrinat et al. [33]، يقيم كل مستخدم ستة معايير للفندق (النظافة، الراحة، الموقع، التسهيلات، فريق العمل، القيمة مقابل السعر)، ويعبَّر عن تفضيلات المستخدم باستخدام تعابير لغوية ضمن سلم من 1 إلى 7 (جيد جداً، جيد، جيد بشكل وسطي، وسطي، سيء بشكل وسطي، سيء، سيء جداً). ثم تُنمذَج قيمة التفضيل لكل معيار ليس باستخدام قيمة عددية وحيدة وإنما باستخدام "قيمة عائمة" (بمجال من القيم بشكل أساسي). يمكن أن نحصل على وزن (تثقيل) كل معيار من قبل المستخدم نفسه، بحيث يتم تمثيل تفضيلات المستخدم الشخصية على المعايير الستة.

تم تقديم خوارزميات قائمة على المنطق العائم لأنظمة الترشيح التعاوني متعددة المعايير في العمل المقدّم من Nilashi et al [35]، بما فيها الاعتماد على المنطق العائم المثقّل Weighted Fuzzy MC-CF والمسافة الإقليدية العائمة Fuzzy Euclidean MC-CF التي تستخدم مقياس تشابه وسطي عائم ومقياس مسافة إقليدي عائم على التوالي، والوسطي العائم Fuzzy Average MC-CF التي تستخدم توقعات عائمة قائمة على المستخدم بشكل مثقّل. قدم Palanivel و Sivakumar [34] مقاربة تجميعية عائمة تقوم بإيجاد معيار تفضيلي باستخدام معامل أعظمي، وبشكل محدد من مؤشرات اهتمام ضمنية مثل الوقت الذي صرفه المستخدم في الاستماع إلى أغنية معينة، أو عدد مرات الولوج إلى أغنية معينة، وتحميل الأغنية. إن استخدام هذه المؤشرات الضمنية يمكن أن يخفف عناء المستخدم للمحافظة على تزويد التقييمات لكل مادة مستهلكة.

2.2.4 المقاربات المعتمدة على النموذج Model–based Approaches

تقوم المقاربات المعتمدة على النموذج ببناء نموذج توقّعي لتقدير التقييمات غير المعروفة عن طريق التعلم من المعطيات المعروفة. تقع العديد من مقاربات أنظمة التوصية متعددة المعايير في هذا التصنيف، مثل التوابع التجميعية البسيطة والنماذج الاحتمالية وطريقة تحليل القيمة الوحيدة MSVD وتوابع التجميع المعقدة SVR.

مقاربة التوابع التجميعية Aggregation Functions Approach

بينما يتم اعتبار التقييم الكلي r_0 كمعيار آخر ضمن قائمة المعايير في المقاربات التجريبية البسيطة (كما وضحنا سابقاً)، نعتبر هذا التقييم على أنه حصيلة تجميع التقييمات متعددة المعايير [8]. بأخذ هذا الافتراض بعين الاعتبار، تقوم هذه المقاربة بإيجاد تابع تجميعي f يمثل العلاقة بين التقييم الكلي والتقييمات الإفرادية:

$$r_0 = f(r_1, \dots, r_k)$$

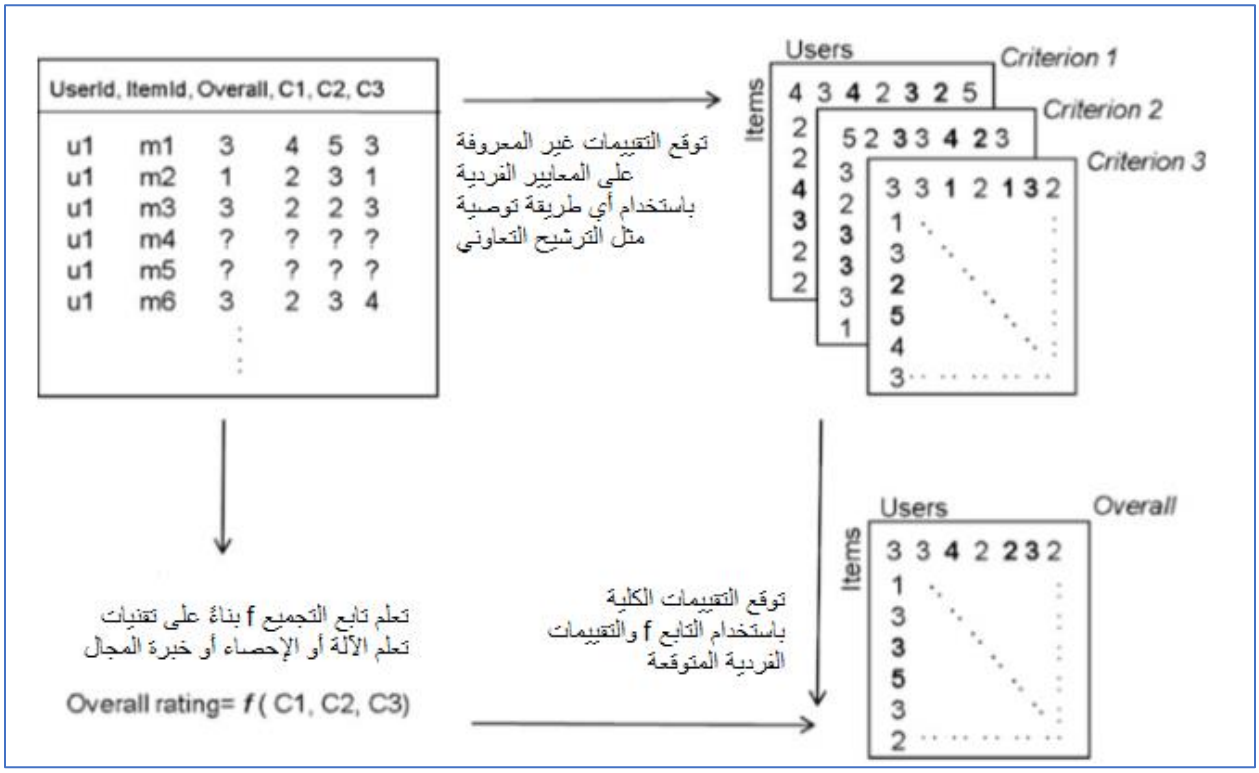
على سبيل المثال، في نظام توصية أفلام، يمكن لمعيار القصة أن يكون ذو أهمية كبيرة، أي أن الأفلام ذات التقييم العالي للقصة سيكون لها تقييم كلي عالي بغض النظر عن التقييمات الأخرى.

تتكون مقاربة التابع التجميعي من ثلاث خطوات، كما هو موضح في الشكل 3. في البداية، يتم توقع k تقييم إفرادي باستخدام أي خوارزمية توصية (مثل الترشيح التعاوني مثلاً). مما يؤدي إلى تقسيم المسألة إلى k مسألة توقع تقييم فردي. ثانياً، يتم اختيار تابع التجميع f بناءً على الخبرة في المجال أو التقنيات الإحصائية، أو تقنيات تعلم الآلة. على سبيل المثال، يمكن للخبرة في المجال أن تقترح تابع وسطي بسيط للتقييمات الإفرادية بناءً على المعرفة والخبرة السابقة في مجال الدراسة. يمكن أيضاً الحصول على تابع التجميع باستخدام التقنيات الإحصائية مثل تقنيات الانحدار الخطي أو اللاخطي، ويمكن أيضاً استخدام تقنيات تعلم الآلة مثل الشبكات العصبونية. في النهاية، يتم حساب التقييم الكلي لكل مادة غير معروفة التقييم بناءً على التقييمات الفردية المتوقعة والتابع f المختار.

بينما يتم تطبيق التقنيات التجريبية المعتمدة على التشابه على مقاربات التوصية المعتمدة على الترشيح التعاوني المعتمد على الجوار فقط، تتمكن مقاربة التابع التجميعي من استخدام أي خوارزمية توصية تقليدية، وذلك بسبب استخدام هذه الخوارزميات في الخطوة الأولى فقط أثناء توقع التقييمات الفردية. كمثال على قابلية ذلك، استخدم [8] الانحدار الخطي والأوزان المتوقعة (أوزان الأهمية لكل معيار) بناءً على التقييمات المعروفة.

لاحظ [8] أيضاً أن تابع التجميع يمكن أن يكون له مجالات مختلفة مثل: الجميع (عندما يتم تعلم تابع التجميع بناءً على مجموعة البيانات كاملةً)، أو معتمدة على المستخدم أو معتمدة على المادة (عندما يتم تعلم تابع تجميع خاص بكل مستخدم أو مادة).

أظهرت النتائج التجريبية من نظام Yahoo! Movies أن استخدام تابع تجميعي يفوق في الأداء استخدام تقنية الترشيح التعاوني ذات التقييم الوحيد (التقييم الكلي) بنسبة 0.3-6.3% بالنسبة للدقة لأعلى N مادة (N=3, 5, 7) [8].



الشكل 3 مقاربة التابع التجميعي (مثال على 3 معايير إفرادية)

تعتمد بعض مقاربات التوصية متعددة المعايير على خوارزميات نمذجة تعتمد على الاحتمالات التي تزداد شهرتها في علوم التنقيب في المعطيات وتعلم الآلة. مثال على هذه المقاربات ما طرحوه .Sahoo et al في ورقتهم البحثية [36] الذي طوروا فيها نموذج المزيج المرن Flexible Mixture Model FMM المطروح من قبل Si وJin [37] لأنظمة التوصية متعددة المعايير.

3.4 تضمين التقييم متعدد المعايير أثناء التوصية Engaging Multi-Criteria Ratings During Recommendation

كما ذكرنا سابقاً، يمكن في أنظمة التوصية متعددة المعايير أن تنمذج منفعة المستخدم تجاه مادة معينة عن طريق تضمين كل من التقييم الكلي والتقييمات الفردية على المعايير الفردية المختلفة، أو أن تقوم بتضمين التقييمات الفردية فقط.

في حال تضمين التقييمات الكلية كجزء من النموذج، تكون عملية التوصية في هذه الحالات مباشرة: بعد توقع كل التقييمات غير المعروفة، يستخدم نظام التوصية التقييم الكلي للمواد لاختيار المواد الأعلى تقييماً لكل مستخدم. أي أن عملية التوصية هي مشابهة للعملية التقليدية التي تعتمد على المعيار الوحيد.

في حال عدم تضمين التقييم الكلي، تصبح عملية التوصية أكثر تعقيداً، وذلك بسبب عدم وضوح بناء ترتيب المواد بشكل سهل. تم توضيح المشكلة القائمة في هذه الأنظمة في الشكل 2. تم طرح العديد من المقاربات في أدبيات أنظمة التوصية الرقمية متعددة المعايير للتعامل مع هذه المشكلة: حاول البعض تصميم ترتيب كلي للمواد والحصول على حل أمثلي عام وحيد لكل مستخدم، بينما حاول آخرون أخذ أحد الترتيبات الجزئية للمواد وإيجاد عدة حلول (أمثلية من نمط Pareto). نسرد في الفقرات التالية الأعمال المنجزة والمتعلقة بأمثَلة تعدد المعايير، وتوصيف العديد من المقاربات التي تم استخدامها في أنظمة التوصية ونناقش الاستخدامات الممكنة للتقييمات متعددة المعايير في عملية التوصية.

4.4 أمثَلة تعدد المعايير Multi-Criteria Optimization

جرى دراسة مسائل الأمثّلة متعددة المعايير بشكل كبير في بحوث العمليات [38]، على الرغم من أنها ليست في سياق أنظمة التوصية. تساعد الأمثَلة متعددة المعايير متخذ القرار على اختيار البديل الأفضل عند حدوث تضارب (متعدد المعايير) بين البدائل. تم استخدام المقاربات التالية غالباً للتوجه إلى دراسة مشاكل الأمثَلة متعددة المعايير، والتي يمكن تطبيقها في أنظمة التوصية، كما تم النقاش في [6]:

- إيجاد الحلول الأمثلية من نمط Pareto.
- تشكيل تركيبة خطية من المعايير المتعددة وإنقاص المسألة إلى مسألة أمثَلة وحيدة المعيار.
- أمثَلة معيار وحيد (الأكثر أهمية) واعتبار المعايير الباقية على أنها قيود.
- أمثَلة معيار وحيد في المرة الواحدة تباعاً، بحيث يتم تحويل الحل الأمثَلي إلى قيود وإعادة الإجرائية لكل المعايير الأخرى.

في أنظمة التوصية متعددة المعايير، يمكن للمادة أن يتم تقييمها بشكل مختلف على المعايير المختلفة، لذلك فإن عملية إيجاد أفضل مادة بشكل عام هي عملية ليست سهلة. نصف لاحقاً عدة مقاربات توصية من التي تم استخدامها في أنظمة التوصية،

حيث تملك جميعها جذوراً في تقنيات الأمثّلة متعددة المعايير، مثل: تحويل مسألة الأمثلة متعددة المعايير إلى مسألة ترتيب وحيدة المعيار، وإيجاد الحلول الأمثلية من نمط Pareto وأخيراً، استخدام المعايير المتعددة كقيود.

1.4.4 تصميم ترتيب كلي لتوصية المواد

هنالك بعض الأعمال في مجال أنظمة التوصية الرقمية التي استخدمت نظريات المنفعة متعددة المعايير في علوم دعم القرار، والتي يمكن وصفها على أنها طريقة لتشكيل تركيبة خطية من المعايير وإيجاد الحل الأمثّل [39]، لنصل في النهاية إلى تخفيض مسألة الأمثّلة متعددة المعايير إلى مسألة ترتيب وحيدة المعيار بسيطة. على سبيل المثال، تقوم المقاربة المقدمة في [39] على ترتيب المواد بالاعتماد على طريقة المنفعة التجميعية UTilités Additive (UTA) المقدمة من .Siskos et al [40]. تهدف خوارزميتهم إلى تقدير الفائدة الكلية U لمادة معينة لكل مستخدم عن طريق إضافة المنافع الهامشية marginal utilities لكل معيار c $(c = 1, \dots, k)$.

$$U = \sum_{c=1}^{k} u_c(R_c) \qquad \text{Eq19}$$

والتي تتبع للقيود التالية:

$$u_c(R_c^{worst}) = 0\ , \forall\ c = 1,2, \dots, k \qquad \text{Eq20}$$

$$\sum_{c=1}^{k} u_c\left(R_c^{best}\right) = u_1\left(R_1^{best}\right) + u_2\left(R_2^{best}\right) + \cdots + u_k\left(R_k^{best}\right) = 1 \qquad \text{Eq21}$$

حيث أن R_c هو التقييم على المعيار c، و$u_c(R_c)$ هو تابع حقيقي غير متناقص (تابع المنفعة الهامشية) لمستخدم معين. باعتبار أن المجال $[R^{worst}, R^{best}]$ هو مجال تقييم المعيار، وأن R_i^{best} و R_i^{worst} هما الحدان الأسوأ والأفضل للمعيار رقم i على الترتيب.

نقوم بسؤال صانع القرار لتزويدنا بتقييمه العام وذلك لتشكيل ترتيب كامل من البدائل (المواد): $i_1 \succ i_2 \succ \cdots \succ i_m$. نعتبر نموذج المنفعة المطوَّر على أنه متوافق مع طريقة محاكمة صانع القرار ليكون $\mathrm{U}(i_1) > U(i_2) > \cdots > U(i_m)$. عند تطوير نموذج المنفعة العام ليتوافق مع هذه المتطلبات، يمكن حصول نوعين من الأخطاء المحتملة وهما: (1) خطأ التقدير بأقل من القيمة (الإبخاس Under Estimation) عندما يقوم النموذج المطور بإسناد البديل إلى ترتيب أقل (أفضل) من البديل المسنَد في الترتيب الكلي (أي أن البديل قد تم تقديره بأقل من قيمته الواقعية من قبل صانع القرار)، و (2) خطأ التقدير بأكبر من القيمة (المغالاة في التقدير Over Estimation) عندما يقوم النموذج بإسناد البديل إلى ترتيب أعلى (أسوأ) من البديل المسند في الترتيب الكلي (أي أن البديل قد تم تقديره بأكثر من قيمته من قبل صانع القرار). يتم اختيار النموذج النهائي عن طريق تصغير مجموع هذين الخطأين. يمكن الحصول على هذا النموذج عن طريق تقنيات البرمجة الخطية مع وجود التقييمات المتوقعة على المعايير المتعددة.

بما أن هذه المقاربة تستخدم معلومات الترتيب مع تقنيات الانحدار العادي Ordinal Regression، يتم استخدام مقياس Kendall's Tau [41] كمقياس ترابط بين متحولين ترتيبيين لمقارنة الترتيب الواقعي مع الترتيب المتوقع.

يطرح [29] طريقة تقوم بحساب المنفعة الكلية U إما عن طريق جمع k منفعة جزئية متوقعة u_k (من خلال خوارزميتهم "التشابه عير المنفعة الجزئية")، أو عن طريق تثقيل التقييمات المتوقعة (في خوارزميتهم "التشابه عير التقييم"). في كلا الحالتين، تُحسب المنفعة الكلية لمادة معينة عن طريق تابع تجميعي على الشكل:

$$U = \sum_{c=1}^{k} u_c = \sum_{c=1}^{k} w_c R_c \qquad \text{Eq22}$$

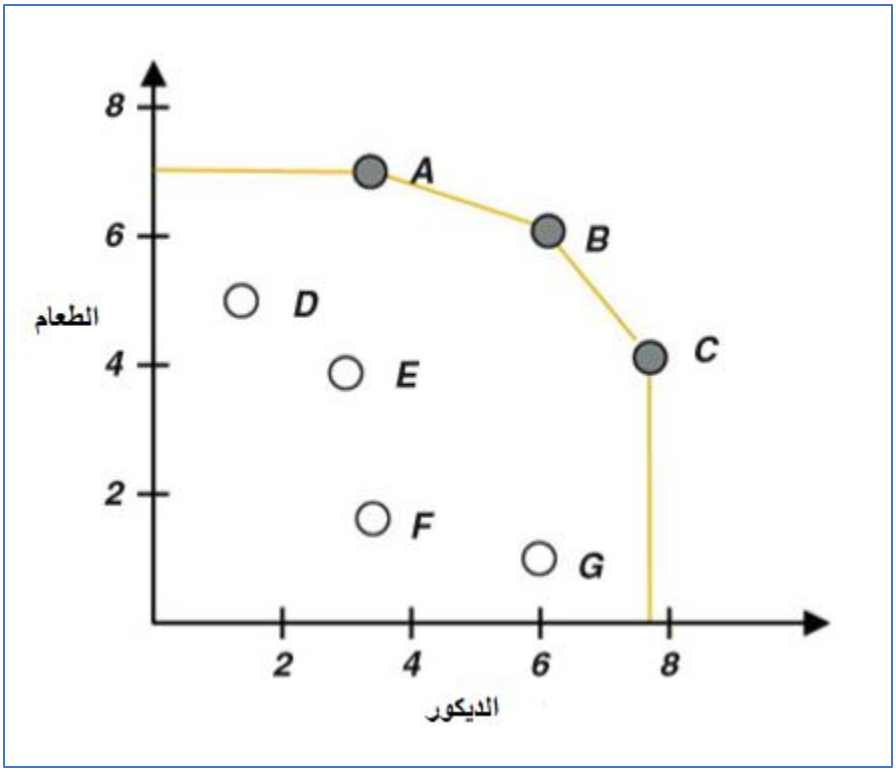
نستخدم هنا التقييمات الفردية على المعايير المتعددة لترتيب المواد المرشحة، بدلاً من توقع التقييمات الكلية بشكل صريح. أخيراً، بعد الحصول على الترتيب الكلي للمواد المرشحة باستخدام أي من التقنيات أعلاه، يتم توصية المواد لكل مستخدم التي توصل المنفعة الكلية إلى أقصى حد.

يستخدم [42] تقييمات كل مستخدم أيضاً لترتيب المواد كقائمة اقتراحات. للقيام بهذا، نقوم بإسناد قيم مثالية لكل مستخدم على كل معيار بمعدل (وسطي) تقييماته السابقة، ونقوم بتوقع تقييم مادة جديدة لمستخدم معين عن طريق حساب المسافة بين القيم المثالية لكل المستخدمين مع القيم المثالية للمستخدم الهدف. ثم عندما يشاهد مستخدم مادةً معينة، يمكن اقتراح أقرب مادة بناءً على التشابهات بين المواد للمستخدم.

2.4.4 إيجاد التوصيات الأمثلية من نمط Pareto

تكتشف هذه المقاربة عدة مواد مناسبة من بين عدد كبير من المواد المرشحة (بدل الوصول إلى حل وحيد عن طريق حل مشكلة الأمثَلة العامة) وذلك عند ربط مواد مختلفة مع عدة معايير متناقضة والترتيب الكلي على المواد غير موجود بشكل مباشر. تُستخدم طريقة تحليل محفظة البيانات Data Envelopment Analysis (DEA)، أو ما يدعى عادة بتحليل الجبهة Frontier Analysis، لقياس فعالية أداء وحدات صنع القرار Decision Making Units (DMU) في بحوث العمليات [43]. تحسب طريقة تحليل محفظة البيانات جبهة الأداء Efficiency Frontier، التي تعرف المواد "الأفضل أداءً" بشكل عام، بأخذ كل المعايير بعين الاعتبار. لا تتطلب هذه الطريقة أمثال للمعايير مقدماً، وتستخدم البرمجة الخطية للوصول إلى المجموعة الأفضل من الأمثال لكل وحدة صنع قرار DMU. بشكل خاص، في سياق أنظمة التوصية متعددة المعايير، بوجود المواد المرشحة التي تكون موجودة للتوصية لمستخدم معين (بما فيها معلومات عن تقييماتها المتوقعة لكل المعايير)، تكون طريقة DEA قادرة على تحديد المجموعة المختصرة للمواد (أي الواقعة على الحدود The Frontier) التي تملك أفضل التقييمات على كل المعايير ضمن جميع المواد المرشحة. يمكن توصية هذه المواد إلى المستخدم بعدها.

بينما لم يتم استخدام طريقة DEA بشكل مباشر في أنظمة التوصية متعددة المعايير، يمكن أن يتم تمثيل مسألة التوصية متعددة المعايير بدون التقييمات الكلية أيضاً على أنها مسألة استعلام عن بيانات في مجال قواعد البيانات باستخدام نفس الطريقة [44]. يستخدم Lee & Teng [44] استعلامات skyline لإيجاد أفضل المطاعم عبر عدة معايير (الطعام والديكور والخدمة والكلفة). كما يظهر الشكل 4، تقوم استعلامات خط المحيط (أو النقاط المسيطرة) skyline بعض النقاط المحيطية (النقاط الأمثلية من نمط Pareto) التي لا يطغى عليها أي نقطة أخرى من عدد كبير من المطاعم المرشحة ضمن فضاء ثنائي الأبعاد (الطعام والديكور). نعتبر هنا، لمستخدم معين، أن النقطة مسيطر عليها، إذا وُجد مادة مرشحة أخرى بحيث يكون لها تقييم أفضل أو مساوٍ لهذه النقطة في كل المعايير.



الشكل 4 مثال على النقاط المسيطرة (المطاعم الأفضل) في فضاء ثنائي الأبعاد

تظهر النتائج التجريبية باستخدام التقييمات متعددة المعايير في استبيان المطاعم في نظام Zagat[4] أن نظام التوصية باستخدام النقاط المسيطرة skyline تساعد على تقليل عدد الخيارات التي يمكن للمستخدم أن يأخذها بعين الاعتبار في استعلاماته. على

[4] https://www.zagat.com/

سبيل المثال، عند استعلام المستخدم عن المطاعم التي تقدم بوفيه مفتوح والتي تقع في مدينة نيويورك مع كلفة لا تتعدى 30 دولاراً، يقترح النظام مطعمَين فقط من أصل 12 مطعماً مرشحاً، وذلك بناءً على التقييمات على أربعة معايير.

على كل حال، يجب توسعة هذا العمل الأولي في عدة مناحي وذلك لأن استعلامات النقاط المسيطرة (أو خط المحيط Skyline) يمكن ألا تتوسع بشكل جيد مع ارتفاع عدد المعايير، مما يؤدي إلى عدد كبير من النقاط المسيطرة مع جهد وكلفة حسابية عالية.

3.4.4 استخدام التقييم متعدد المعايير كمعاملات تصفية للاقتراحات

بشكل مشابه لكيفية استعمال مواصفات المحتوى كمرشحات للتوصية في أنظمة التوصية [45] [46]، يمكن استخدام التقييمات متعددة المعايير لنفس الأهداف أيضاً. على سبيل المثال، يمكن للمستخدم أن يحدد الأفلام ذات القصة الممتازة يجب أن يتم توصيتها له في وقت معين، بغض النظر عن المعايير الأخرى، مثل المؤثرات البصرية، ومن ثم، يتم توصية الأفلام ذات التقييم العالي المتوقع على معيار القصة (لنقل أكبر من 9 من 10).

يمكن أن نصف هذه الطريقة بكلمات أخرى، يمكن أن ننقص أبعاد مسألة الأمثلة متعددة الأبعاد عن طريق تحويل المعايير إلى قيود (مرشحات Filters). يمكن اعتبار هذه المقاربة على أنها مشابهة لمقاربة الاعتماد على المحتوى Content-Based [47] أو المدركة للسياق Context-aware [45] [46]. على كل حال، فهي مختلفة بعض الشيء عنهم، لأن الترشيح يتم ليس بالاعتماد على واصفات المحتوى الموضوعية (مثل طول الفيلم أكبر من 120 دقيقة) أو أبعاد السياق الإضافية (الوقت الحالي هو عطلة نهاية الأسبوع)، وإنما على معايير التقييم الشخصية (مثل تقييم القصة أكبر من 9)، أي القيمة المتوقعة التي تعتمد بشكل كبير على ذوق وتفضيلات المستخدم.

5.4 مناقشة الأعمال المستقبلية

تمثل أنظمة التوصية منطقة بحث حية ومتغيرة باستمرار. عبر التطويرات الحديثة المهمة، بدأت أنظمة التوصية الرقمية بتبني التقييمات متعددة المعايير المقدمة من المستخدمين، ونقوم في هذا القسم باستعراض الخوارزميات والتقنيات لأنظمة التوصية متعددة المعايير. لم يتم دراسة هذه الأنظمة الحديثة نسبياً بشكلٍ وافٍ، ونقدم في هذا القسم بعض التحديات واتجاهات البحث المستقبلية لهذا التصنيف من أنظمة التوصية.

1.5.4 تطوير مقاربات جديدة للتقييمات متعددة المعايير

1.1.5.4 نمذجة التقييمات متعددة المعايير

عادةً ما يتم التعبير عن تفضيلات المستخدم في أنظمة التوصية (بما فيها الأنظمة متعددة المعايير) باستخدام تقييمات رقمية بسيطة. بدأت الأعمال الحديثة باستكشاف مقاربات بديلة لتمثيل وتجميع تقييمات المستخدم (مثل استخدام تقنيات المنطق العائم [48] [49])، إلى جانب ذلك، يمكن نمذجة التقييمات بطريقة ملونة (أو تعتمد على الأسماء) (مثل أخذ المواصفات الدلالية للتقييمات الدلالية [50]). إن دراسة نمذجة تفضيلات المستخدم، وخاصة في سياق متعدد المعايير أكثر تعقيداً، تمثل اتجاه شيق في البحث مستقبلاً.

2.1.5.4 التقسيم والمعالجة السابقة للبيانات بشكل ذكي

إنه لمن المعروف أن أنظمة التوصية تعاني من مشكلة تبعثر البيانات Data Sparsity. أحد المقاربات المحتملة لتخفيف هذه المشكلة هي بتطبيق تقسم للمعطيات بشكل ذكي أو عنقدة لهذه المعطيات، حيث يتم تجاهل الأبعاد (المعايير) غير المفيدة، ويتم دمج المستخدمين المتشابهين (أو المواد المتشابهة)، وتُحسب التوصيات الناتجة عن طريق أخذ هذا التجميع بعين الاعتبار. في أدبيات التنقيب في المعطيات، كان هناك بحثاً كبيراً في كيفية تقسيم المستخدمين [51]. أيضاً، في أنظمة التوصية متعددة المعايير، تم استخدام العديد من إجرائيات عنقدة المستخدمين المخصصة كجزء من خوارزميات التوصية المقترحة (مثل [52] [53]). استكشف بعض الباحثون أيضاً العديد من تقنيات استخلاص الميزات Feature Extraction لتحديد المعايير الأفضل لاستخدامها في الأنظمة متعددة المعايير [54]. على أي حال، نحتاج العديد من الدراسات الإضافية لفحص مقاربات تقسيم ومعالجة البيانات السابقة Data Pre-processing لأنظمة التوصية متعددة المعايير بطريقة دلالية.

3.1.5.4 توقع التفضيلات المشابهة

يمكن دراسة مسألة التوصية متعددة المعايير عن طريق تمثيلها كتوقع (تقدير) التفضيلات المتشابهة للمستخدم، على النقيض لقيم التقييم المطلقة. كان هناك بعض الأبحاث حول تشكيل الترتيب الصحيح المتشابه للمواد باستخدام تقنيات تعتمد على الترتيب.

على سبيل المثال، قام Freund [55] بتطوير خوارزمية RankBoost بالاعتماد على الطريقة المعروفة AdaBoost، وفي إعدادات متعددة المعايير، يمكن لمثل هذه الخوارزميات أن يتم اعتمادها لتجميع معايير التقييم المتشابهة لمستخدم معين. بالتحديد، تم الاعتماد على هذه المقاربة في نظام DIVA [56].

4.1.5.4 بناء معايير أداء المادة

نحتاج للمزيد من البحث حول اختيار أو تشكيل المجموعة الأفضل من المعايير لتقييم المواد. على سبيل المثال، تطلب العديد من أنظمة التوصية متعددة المعايير من المستخدم أن يقوم بتقييم المادة على عدة معايير بمستوى واحد (القصة والمؤثرات الخاصة للفيلم على سبيل المثال). يمكن أن نقوم بتقسيم المستوى الوحيد من المعايير إلى عدة معايير فرعية، بحيث يمكن وجود عدد مستويات حسب المسألة. على سبيل المثال، في نظام توصية الأفلام، يمكن تقسم معيار المؤثرات الخاصة إلى مؤثرات صوتية ومؤثرات بصرية. يمكن لهذه المعلومات الإضافية الناتجة عن المستويات الإضافية للمعايير أن تساعد في فهم تفضيلات المستخدم وفي بعض التقنيات مثل الإجرائية الهرمية التحليلية Analytical Hierarchy Process AHP التي يمكن استعمالها لدراسة هرمية المعايير [57]، كما طرح Schmitt et al. [58] في نظامهم. عند أخذ المزيد من المعايير لكل مادة، يجب أيضاً أن نقوم بفحص الترابط بعناية بين هذه المعايير لأن اختيار المعايير يمكن أو يؤثر بشكل كبير في جودة التوصية. أبعد من ذلك، من المهم الحصول على عائلة متسقة من المعايير لنظام توصية، والتي تعني أن المعايير هي كاملة الترتيب وشاملة وغير زائدة عن الحاجة. في الخلاصة، إن بناء مجموعة من المعايير لمسألة توصية معينة هو مجال مهم وممتع للأبحاث المستقبلية.

5.1.5.4 دمج المعلومات المتعلقة بالمجال

إن العديد من أنظمة التوصية متعددة المعايير هي مصممة بدون استثمار المعرفة الخاصة بالمجال المدروس. على سبيل المثال، يمكن لفهم المعلومات الإضافية لخصائص الفندق (كالنظافة والموقع والخدمات... الخ)، ولكن أيضاً فهم خصائص الأشخاص المسافرين (مثل المسافرين للعمل أو المسافرين ذوي الخبرة في السفر أو المسافرين لقضاء عطلة شهر العسل أو عطلة رومانسية أو المسافرين لأخذ عطلة الربيع... الخ) يمكن أن يزود فوائد كبيرة في تصميم خوارزميات توصية أفضل. قامت العديد من الدراسات بالبدء باكتشاف النماذج التي تتضمن المعلومات الخاصة بالمجال المدروس ضمن أنظمة التوصية متعددة المعايير [59]، ولكن هناك العديد من الفرص الإضافية في هذا مجال هذا البحث.

2.5.4 توسعة التقنيات الحالية للنظم متعددة المعايير

1.2.5.4 إعادة استخدام تقنيات التوصية وحيدة المعيار

تم تطوير عدد كبير جداً من تقنيات التوصية لأنظمة التوصية وحيدة المعيار خلال آخر 15–20 عاماً، وتم توسعة بعضهم إلى أنظمة التقييم متعددة المعايير. على سبيل المثال، يمكن لتقنيات الترشيح التعاوني المعتمدة على الجوار أن تأخذ بعين الاعتبار التقييمات متعددة المعايير باستخدام عدد كبير من خيارات التصميم كما اقترح Manouselis وCostopoulou [29] (كما تم النقاش في المقاربات التجريبية Heuristic Approaches).

ناقشنا في هذا القسم العديد من اتجاهات البحث المستقبلية المحتملة للأنظمة متعددة المعايير التي يمكن أن تكون ممتعة ومهمة في مجتمع أنظمة التوصية. لم يكن الهدف من هذه اللائحة أن تكون شاملة، ولكن نؤمن أن البحث في هذا المجال ما زال في مراحله البدائية، وهناك العديد من المجالات الإضافية التي يمكن استكشافها للتقدم في أنظمة التوصية متعددة المعايير.

6.4 محرّكات توصية مشابهة

تفتقر الشبكة العنكبوتية لمحركات التوصية غير المرتبطة بالسياق (أي التي تكون مستقلة عن النظام الأساسي)، ولكن هناك بعض الخيارات الأخرى التي تعمل وفق الخوارزميات التقليدية وتقدم خيارات تعدد المعايير حسب الطلب.

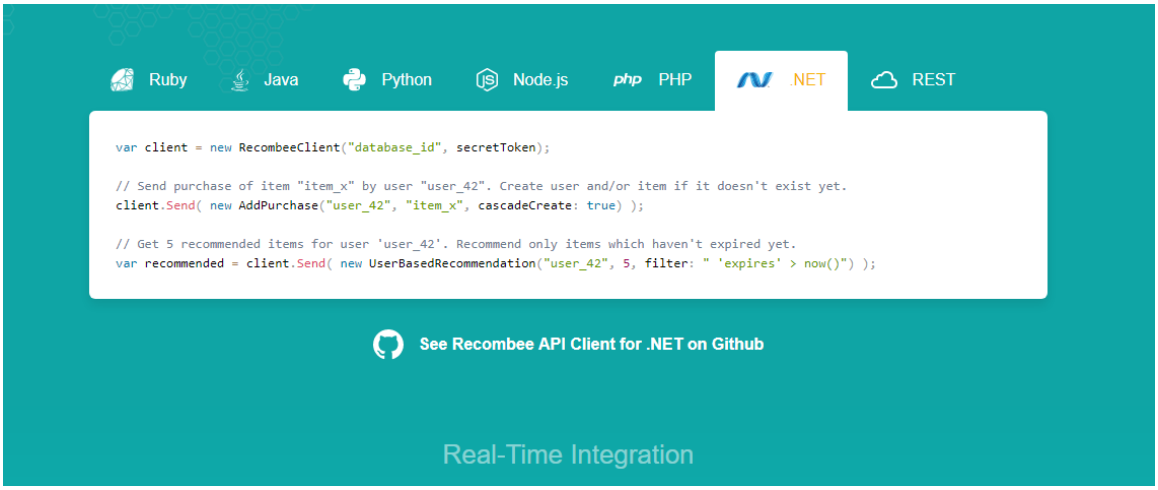
Recombee 1.6.4

يقدّم محرّك التوصيّة الرقميّ Recombee خدمة التوصيّة في الوقت الحالي عبر واجهة برمجة تطبيقات API. تقوم بإعادة التّوصيات خلال أقلّ من 200 م.ث. وهو زمن استجابة ممتاز، كما أنّه قادرٌ على تقديم هذه الخدمة مهما كان الضغط كبيراً على موقع الزّبون حيثُ أنّه قادرٌ على تقديمِ أكثر من 500 توصية في الثّانية الواحدة.

يتيح المحرّك لزبائنه استطلاع مؤشّرات الحساسيّة الخاصّة بهم كدقّة التقييمات المقدّمة من الموقع، وغيرها من المؤشّرات متى ما أرادوا وذلك من خلال واجهة تفاعليّة سهلة الاستخدام. ويتميّز هذا المحرّك بأنّه مستقلٌّ عن مجالِ التّوصية حيث يمكن استخدامه لتوصية أفلام، مقالات، كتب، أو أغاني وغيرها...

يُقدّم هذا المحرّك خدماته مجاناً حتى 100 ألف توصية شهريّاً، ثمّ تتحوّل هذه الخدمة إلى خدمة مدفوعة مقابل 100$ شهريّاً.

يمكن استخدام هذه الخدمة بسهولة كما هو مبيّن في الشّكل التّالي والمأخوذ من الموقع كمثال عن كيفيّة الاستخدام:



الشكل 5 مثال يوضح حالة استخدام لخدمة محرك التوصيّة Recombee

Suggest Grid 2.6.4

يقومُ هذا المحرّك بتحليل بيانات المستخدمين المقدّمة من الزّبائن، وبالتّالي تقديم توصيات مناسبة للمستخدمين حيث تتلخّص خدماته بما يلي:

- توصية أغراض للمستخدمين.
- توصية أغراض لمجموعات من المستخدمين.
- إيجاد مستخدمين محتملين يفضّلون أغراض معيّنة.
- إيجاد مستخدمين مشابهين لمستخدم معيّن.
- إيجاد أغراض مشابهة لغرض معيّن.

كما أنّه يتيح لزبائنه لوحة تحكّم، تمكّنهم من استطلاع مؤشّرات الدقّة الخاصّة ببياناتهم، والتحكّم ببعض الخيارات المتاحة لهم من خلال واجهة تفاعليّة.

يمكن استخدام هذه الخدمة بسهولة كما هو مبيّن في الشكل التالي والمأخوذ من الموقع كمثال عن كيفيّة الاستخدام:

```
var recommendedItemsResponse = suggestGridClient.Recommendation.GetRecommendedItems(new GetRecommendedI
{
    Type = "views",
    UserId = "2",
    Size =   2
});

var recommendedItems = recommendedItemsResponse.Items; // get items
```

الشكل 6 مثال يوضح حالة استخدام لخدمة محرك التوصيّة SuggestGrid

3.6.4 خوارزمية Yong Z.

طرح Yong Z. [60] خوارزميته الجديدة التي تقوم على فكرة استخدام أي تركيبة من المعايير ضمن السياق المدروس لتقوم بتوقع التقييمات متعددة المعايير بشكل أفضل. تبين النتائج التجريبية بناءً على معطيات من TripAdvisor وYahooMovies أن هذه الطريقة تقوم بتحسين أداء التقييمات متعددة المعايير.

قدم الكاتب في هذه الخوارزمية مقاربتي CCA (Criteria Chains: Aggregation Model) و CCC (Criteria Chains: Contextual Model) التان تستخدمان ربح المعلومات information gain لإنتاج ترتيب المعايير الأفضل. بشكل خاص، تسمح مقاربة CCC بمشاهدة التقييمات المتوقعة على أنها أوضاع ضمن السياق، وتكون قادرة على التغلب على مقاربة CCA التي تقوم بتجميع التقييمات متعددة المعايير ضمن نماذج تعتمد على الانحدار الخطي SVR.

7.4 خلاصة

في هذا الفصل، هدفنا إلى تقديم لمحة عن أنظمة التوصية متعددة المعايير. بشكل خاص، ركزنا على أنظمة التوصية المعتمدة على التقييمات متعددة المعايير، أي تلك التقنيات التي تقوم بتزويد الاقتراحات عبر نمذجة تفضيلات المستخدم تجاه مادة معينة كشعاع من التقييمات عبر المعايير المتعددة. قمنا بمراجعة التقنيات الحالية التي تستخدم التقييمات متعددة المعايير لحساب التقييمات المتوقعة وتوليد الاقتراحات، وناقشنا مواضع المناقشة والتحديات المستقبلية لهذا الصنف من أنظمة التوصية.

توفر هذه الدراسة نظرة منهجية لأنظمة التوصية متعددة المعايير، وخارطة طريق للأعمال المتعلقة، ونقاشاً لعدد من اتجاهات البحث المستقبلية. على أي حال، نؤمن أن هذا الصنف من أنظمة التوصية ما يزال في مرحله البدائية من التطوير، ونحتاج الكثير من الأبحاث لفك الإمكانات الكاملة لأنظمة التوصية متعددة المعايير.

الفصل الرابع: تصميم الحل المقترح والنتائج

1.5 مقدمة

تعتمد المقاربة المقترحة على التركيز على فكرة بناء عائلة متسقة من المعايير، حيث نذكر بتعريف العائلة المتسقة من المعايير على أنها (حسب 2.5.3 عائلة المعايير):

1. **كاملة الترتيب Monotonic**. نقول عن مجموعة من المعايير أنها كاملة الترتيب فقط إذا تحقق الشرط: من أجل كل ثنائية بدائل i_1, i_2، وعندما يكون $g_{c1}(i_1) > g_{c1}(i_2)$ من أجل معيار معين c_1، و $g_c(i_1) = g_c(i_2)$ من أجل كل المعايير $c \neq c_1$، يمكن عندها اعتبار البديل i_1 أفضل من البديل i_2.
2. **شاملة Exhaustive**. نقول عن مجموعة من المعايير أنها شاملة فقط إذا تحقق الشرط: من أجل كل ثنائية بدائل i_1, i_2، وعندما يكون $g_{c1}(i_1) = g_{c1}(i_2)$ من أجل كل معيار c، يمكن عندها اعتبار البديلين i_1 و i_2 متكافئين.
3. **غير زائدة عن الحاجة Non-redundant**. نقول عن مجموعة من المعايير أنها غير زائدة عن الحاجة فقط إذا أدّى حذف أي معيار من هذه المجموعة إلى مخالفة الخاصتين السابقتين.

تقوم الفكرة الأساسية ضمن النموذج المقترح على صعوبة تشكيل عائلة معايير غير مرتبطة ببعضها البعض (أي مستقلة بشكل كامل) أو حسب النقاط السابقة، غير زائدة عن الحاجة، وخصوصاً مع وجود شرط الشمولية الذي يفرض اكتمال الصورة لتقييم المستخدم للمادة بشكل صحيح وكامل.

إن عدم استقلالية هذه المعايير يمكننا من الاعتماد على ارتباطها أثناء اكتشاف التوقعات على مستوى كل معيار (باستخدام تقنيات التوصية وحيدة المعيار التقليدية مثل الترشيح التعاوني).

2.5 تذكرة بإشكالية البحث

تعتمد العديد من أنظمة التوصية الرقمية على معيار وحيد يعبر عن تقييم المستخدم للمادة. إن الاعتماد على تقييمات متعددة للمادة يعطي معلومات أكثر عن تفضيل المستخدم للمادة وهو ما اعتمده موقع Trip Advisor [5] الشهير في تصميم نظام التقييم لديه. يوضح المثال التالي مشكلة الاعتماد على تقييم وحيد للفنادق.

[5] https://www.tripadvisor.com/

المستخدم	المريديان	الشيراتون	الشام	أمية
محمد	2، 3، 5، 2، 3	5، 2، 3، 2، 3	2، 3، 5، 2، 3	2، 5، 3، 2، 3
علي	5، 2، 2، 3، 3	2، 5، 2، 3، 3	5، 3، 2، 2، 3	3، 2، 2، 5، 3
مجد	2، 3، 5، 2، 3	2، 3، 5، 2، 3	2، 3، 5، 2، 3	2، 3، 5، 2، 3

تمثل الأرقام الموجودة تقييم المستخدم للمعايير التالية على الترتيب (الخدمة، المكان، النظافة، الجو العام) ويمثل الرقم الأخير الوسطي الحسابي للمعايير الأربعة.

تبرز المشكلة عند الاعتماد على تقييم وحيد (القيمة الخامسة في خلايا جدول التقييمات التي تمثل التقييم الكلي) في اكتشاف النظام لتشابه المستخدم الهدف وهو محمد لكل من علي ومجد على التساوي على اعتبار أن التقييمات مشابهة بشكل كامل بين المستخدمين الاثنين لمحمد.

لدى الاعتماد على التقييمات التفصيلية تكتشف عدم موافقة علي لمحمد في آراءه التفصيلية بينما نرى أن مجد مشابه بشكل تام لمحمد.

إن الاعتماد على المعايير التفصيلية يزيد من دقة نظام التوصية الرقمي. ندعو هذه المقاربة **بأنظمة التوصية متعددة المعايير Multi Criteria Recommender Systems**.

سنعرض فيما يلي مقاربتنا لنموذج مقترح يساهم في بناء نظام توصية متعدد المعايير بالاستفادة من فكرة صعوبة بناء عائلة متسقة من المعايير.

3.5 بناء عائلة المعايير

نعتمد في بناء عائلة المعايير على رأي الخبير في المجال لتصميم هذه العائلة. يعتمد تصميم هذه المجموعة على المجال المدروس مثل:

- نظام توصية مطاعم: جودة الطعام والديكور والموقع، أو الأسعار والخدمة والطعام والتقييم الكلي.
- نظام توصية أفلام: القصة والمؤثرات البصرية والإخراج، أو جودة التمثيل والموسيقى التصويرية والقصة.
- نظام توصية كتب: القصة والتشويق وجودة تقسيم الفصول.

نلاحظ أن كل مجال يحتمل وجود عائلة كبيرة جداً من المعايير التي يمكن تشكيل أي عائلة جزئية منها. إن صعوبة الشرط "عدم الزيادة عن الحاجة" يجعل انتقاء هذه العائلة صعباً بوجود شرط شمولية العائلة لشروط المسألة بحيث تغطي كافة المعايير الممكنة للمسألة التي تجعل مسألة التقييم التقليدية جيدة لمستخدمي النظام.

إن شمولية عائلة المعايير توصل بشكل غير مباشر إلى زيادة المعايير عن الحاجة مما يؤدي إلى تعلق المعايير ببعضها البعض مثل تعلق معيار إخراج الفيلم بالمؤثرات البصرية والسمعية أو تعلق الخدمة في المطعم بالأسعار (وذلك بسبب التقييم السياحي للمطاعم). يؤدي تعلق المعايير ببعضها البعض إلى وجود ترتيب لهذه المعايير يعبر عن تعلقها ببعضها البعض. نسمي عائلة المعايير في المسألة المطروحة بالمجموعة $\mathbb{C}$ حيث:

$$\mathbb{C} = \{c_1, c_2, \dots, c_k\}$$

يمكن ترتيب مجموعة المعايير السابقة بـ $k!$ طريقة مختلفة (حسب قانون التراتيب).

4.5 مجموعة البيانات المستعملة Trip Advisor Dataset

نعتمد في بحثنا على إجراء الاختبارات باستخدام مجموعة البيانات المقدمة من موقع Trip Advisor الشهير لتقييم الفنادق والمطاعم والوجهات السياحية. يوضح الشكل 7 مثالاً عن التقييمات متعددة المعايير في موقع Trip Advisor، حيث يمكن للمستخدمين تقييم الفندق على 7 معايير مختلفة للفندق وهي:

- القيمة مقارنة بالمال المدفوع Value.
- جودة الغرف Rooms.
- ملاءمة موقع الفندق Location.
- نظافة الفندق Cleanliness.
- تجربة تسجيل الدخول للفندق Check in / Front Desk.
- جودة الخدمة المقدمة Service.
- جودة خدمة الأعمال Business Service.

الشكل 7 مثال عن التقييمات متعددة المعايير في موقع Trip Advisor

تم استخراج هذه المعطيات من قبل .Jannach, et al [48] باستخدام إجرائية زحف عبر الوب Web Crawling التي قامت بتجميع تقييمات المستخدمين للفنادق الواقعة في 14 وجهة عالمية مثل لندن ونيويورك وسنغافورة... الخ. يوجد في هذه المجموعة 22,130 تقييماً من قبل 1,502 مستخدم لـ 12,773 فندقاً.

5.5 الخوارزمية المقترحة: سلسة المعايير Criteria Chaining

تعتمد الخوارزمية المقترحة بشكل أساسي على تضمين تعدد المعايير أثناء التوقع وليس التوصية، وذلك عن طريق بناء نموذج يقوم باستخراج التوصيات الفردية وثم تجميعها إلى تقييم كلي.

لنمذجة الخوارزمية رياضياً نعبر عن مسألة التوصية باستخدام التابع:

$$R: Users \times Items \rightarrow R(u,i) \qquad \text{Eq23}$$

بحيث يكتب التابع $R(u,i)$ بدلالة تقييمات المستخدمين الموجودين في جوار المستخدم u ونعبر عن هذا الجوار بـ $\aleph_u$، وبهذا يكون:

$$R(u,i) \sim R(u',i) \quad : \quad u' \in \aleph_u \qquad \text{Eq24}$$

نقصد بـ $R(u,i)$ هنا التقييم المتوقع للمستخدم u للمادة i بينما نقصد بـ $R(u',i)$ بالتقييم الحقيقي المعروف للمستخدم u' للمادة i.

بالواقع لا تتدخل الخوارزمية المقترحة بحساب جوار المستخدم الذي يمكن استخراجه عن طريق خوارزميات تتعلق بطريقة الترشيح التعاوني والتي تعتمد بشكل أساسي على حساب التشابهات بين المستخدم والمستخدمين الآخرين عن طريق التقييمات المعروفة للمواد [28].

ما تطرحه الخوارزمية الجديدة "سلسة المعايير" هو تغيير معادلة حساب التقييم المتوقع لتأخذ بعين الاعتبار الفكرة المطروحة في الفقرة 3.5 بناء عائلة المعايير.

يمكن التعبير عن المعادلة 24 بشكل آخر بحيث يمكن تضمين التقييمات الإفرادية $r_{u,i,c}$ التي تعبر عن تقييم المستخدم u للمادة i على المعيار c.

$$R(u,i) = f\left(r_{u',i,c}\right) \quad : \quad u' \in \aleph_u, c \in C \qquad \text{Eq25}$$

حيث تعبر $\mathbb{C}$ عن عائلة المعايير المدروسة.

تصبح المعادلة 25 على الشكل:

$$R(u,i) = f\left(r_{u',i,c}, r'_{u,i,c'}\right) : \ u' \in \aleph_u, c \in \mathbb{C}, c' \in \mathbb{C}'_c \qquad \text{Eq26}$$

حيث تعبر $\mathbb{C}'_c$ عن مجموعة المعايير التي تسبق المعيار الهدف c التي يتم توقع التقييم عنده.

تحاول خوارزميتنا المعتمدة أن تكامل الاعتماد بين التقييمات متعددة المعايير أو بين المعايير ذاتها ضمن إجرائية التوقع. نقوم بتعلم هذه الخبرة من خوارزمية سلاسل المصنّف Classifier Chains [61] والتي تمثل إحدى المقاربات المستخدمة ضمن التصنيف متعدد العلامات Multi Label Classification (MLC). تكمن الفكرة في خوارزمية سلاسل المصنف أنها قادرة على تعلم التوقعات الثنائية (0 أو 1) لكل وسم (علامة Label) على شكل سلسلة ضمن عملية التصنيف. ففي البداية، تستخدم الخوارزمية الميزات الموجودة في النظام لتستخرج التوقعات للوسم الأول. ثم يتم استخدام هذا التوقع كميزة إضافية للمعطيات لتوقع الوسم الثاني. يتم إضافة توقعات الأوسمة كميزات إضافية لتوقع الوسم التالي، حتى يتم توقع كل الأوسمة. إن خوارزمية سلسلة المصنف قادرة على التفوق على العديد من خوارزميات التصنيف متعددة العلامات أو الأوسمة حيث أنها تأخذ الاعتماديات بين الأوسمة بعين الاعتبار.

بالاعتماد على المعادلة 25، وبالاعتماد على الانحدار الخطي Linear Regression، يمكن كتابة التقييم الكلي على الشكل:

$$R_0 = w_1 * R_1 + w_2 * R_2 + \cdots + w_k * R_k + t \qquad \text{Eq27}$$

يتم حساب التقييم المتوقع R_i باستخدام التقييمات السابقة المتوقعة حيث يتم حساب التقييم على المعيار الأول R_1 طريقة الترشيح التعاوني التقليدية على المعيار الأول، ثم يتم حساب التقييم على المعيار الثاني R_2 باستخدام تقييمات المستخدمين الآخرين (جوار المستخدم ضمن طريقة الترشيح التعاوني) والتقييم المتوقع على المعيار الأول R_1، وهكذا دواليك، يتم حساب التقييم المتوقع على المعيار i باستخدام تقييمات المستخدمين الآخرين (الموجودين في الجوار) على المعيار i والتقييمات المتوقعة على المعايير $1 \ldots i-1$ لنفس المادة. يمكن كتابة معادلة حساب التقييم R_i كالتالي:

$$R_i = \overline{R_u} + \frac{\sum_{v \in \mathcal{N}_i(u)} w_{uv} \left(R_{vi} - \overline{R_v}\right)}{\sum_{v \in \mathcal{N}_i(u)} |w_{uv}|} + w_1 * R_1 + w_2 * R_2 + \cdots + w_{i-1} * R_{i-1} \qquad \text{Eq28}$$

يمثل الحد الأول $\overline{R_u} + \frac{\sum_{v \in \mathcal{N}_i(u)} w_{uv} \left(R_{vi} - \overline{R_v}\right)}{\sum_{v \in \mathcal{N}_i(u)} |w_{uv}|}$ قيمة التقييم اعتماداً على الترشيح التعاوني.

تعبر الأمثال w_{uv} عن مقدار الترابط Correlation بين المستخدمين u, v، بينما تعبر الأمثال w_j عن الترابط بين التقييمات الموجودة على المعيار الحالي i والمعيار j. يمكن حساب قيمة الترابط باستخدام مقياس Pearson[6] [62].

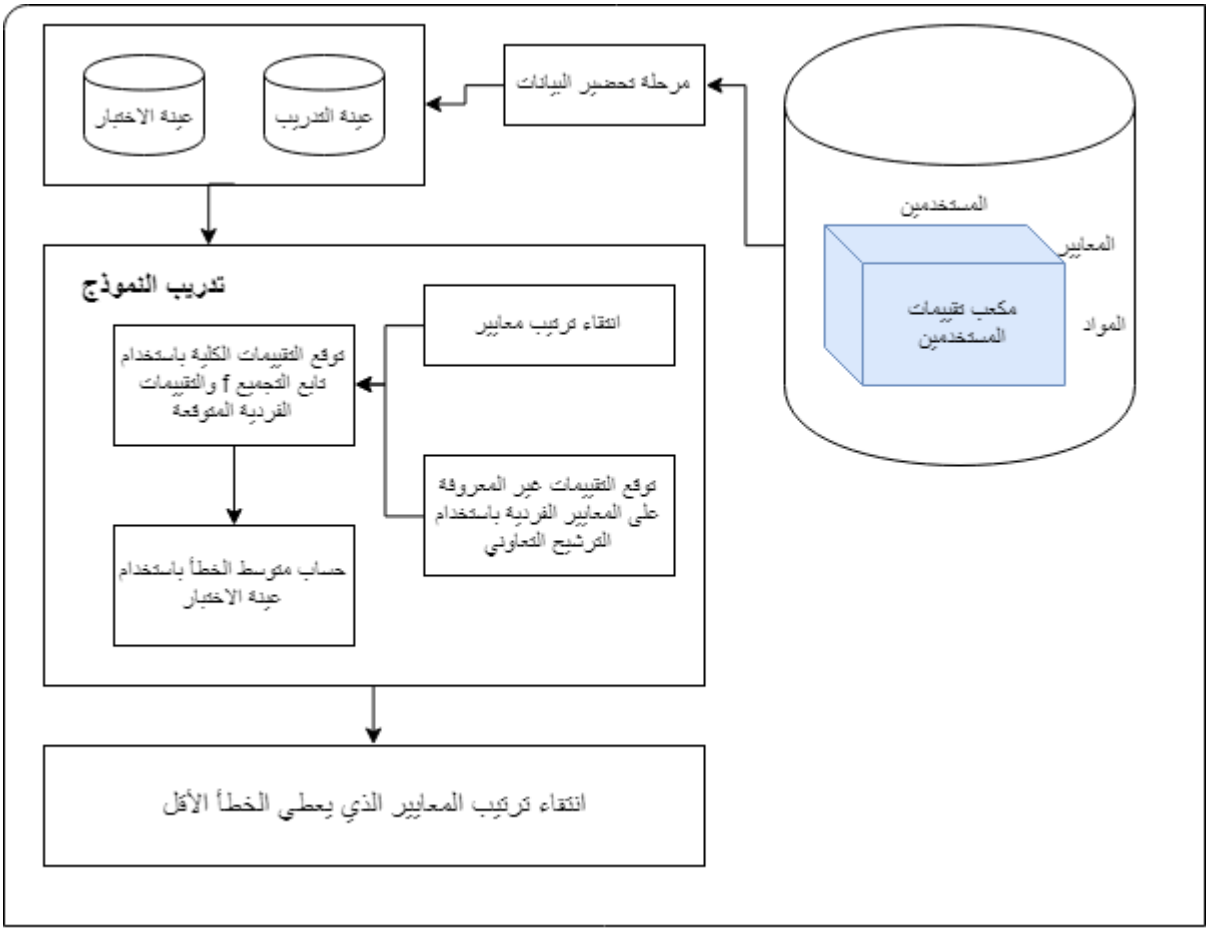
كما سبق وذكرنا سابقاً وجود $k!$ ترتيب ممكن لـ k معيار، بالتالي يمكن ترتيب المعايير لاستخدامها في الخوارزمية ضمن الجزء الخاص بسلسلة المعايير ضمن المعادلة 28.

مثال: ضمن مجموعة المعطيات المدروسة Trip Advisor يمكن ترتيب المعايير ضمن $7! = 5040$ طريقة.

تعمل الخوارزمية المقترحة على دراسة كافة الترتيبات الممكنة للمعايير وحساب الخطأ التربيعي الأصغري MAE الناتج في كل ترتيب (عن طريق اختبار النتائج على عينة الاختبار)، ومن ثم انتقاء الترتيب الذي يعطي الخطأ الأقل.

تقوم الخوارزمية ببناء نموذج تدريب يعمل ضمن مرحلتين أساسيتين: مرحلة غير متصلة Offline ومرحلة متصلة Online. يتم تدريب النموذج خلال المرحلة غير المتصلة بينما يتم تطبيقه لإعطاء النتائج خلال المرحلة المتصلة. يبين الشكل 8 مراحل عمل الخوارزمية المقترحة.

[6] راجع الملحق 1 للاستزادة عن الترشيح التعاوني وعن طريقة حساب الترابط بطريقة Pearson وبطرق أخرى مثل الزوايا cos.



الشكل 8 خوارزمية سلسلة المعايير

تعمل الخوارزمية على تحضير البيانات وتقسيمها إلى عينتين هما عينة التدريب وعينة الاختبار بحيث يتم اقتسام المعطيات إلى 80% للتدريب و20% للاختبار.

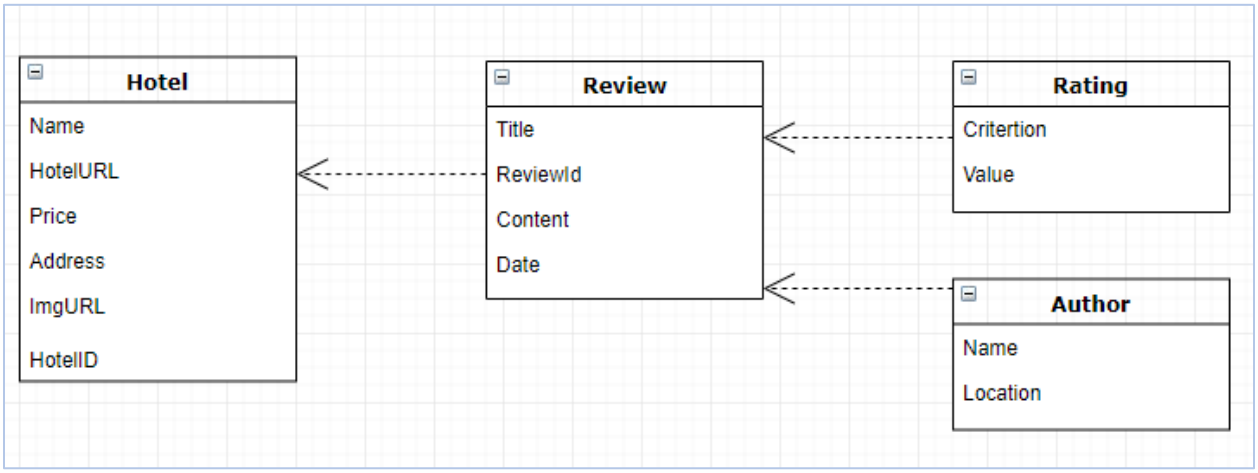
6.5 تطبيق خوارزمية سلسلة المعايير على بيانات موقع Trip Advisor

جرى تطبيق الخوارزمية المقترحة على بيانات Trip Advisor ضمن عدة مراحل وهي:

1- المرحلة الأولى: تحضير البيانات.

تتوفر البيانات ضمن عدة ملفات من نوع json بحيث تكون يمثل كل ملف التقييمات الخاصة بكل فندق. يحوي الملف على قائمة بالتقييمات بحيث يحوي كل تقييم على: التقييمات على المعايير المتعددة واسم وموقع المستخدم وعنوان ورقم المراجعة ومحتوى المراجعة النصي وتاريخها.

يقوم محضر البيانات Data Preprocessor بقراءة هذه الملفات واستيرادها إلى بنية معطيات مقابلة موضحة في الشكل 9 بنية المعطيات الممثلة لتقييمات بيانات Trip Advisor.



الشكل 9 بنية المعطيات الممثلة لتقييمات بيانات Trip Advisor

بعد تحضير البيانات نقوم بتجزئتها إلى عينتين هما عينة التدريب (80%) وعينة الاختبار (20%). تم تجزئة العينتين عبر اقتسام المراجعات وليس الفنادق وذلك للحصول على عينتين عادلتين للتدريب والاختبار لنكون قادرين على الاختبار بشكل عادل.

2- المرحلة الثانية: تدريب واختبار النموذج.

نقوم بالاعتماد على تطبيق خوارزمية الترشيح التعاوني (باستخدام مقياس Pearson) لاستنتاج التقييمات الإفرادية، كما نستخدم مقياس Pearson في استخراج أمثال المعادلة الخطية الخاصة بسلسلة المعايير حيث يعبر هذا المعيار عن مقدار الترابط بين المعيارين المدروسين باستخدام تقييمات المستخدمين على المعيارين.

نقوم باختبار كافة تركيبات المعايير الممكنة ($k!$ تركيبة ممكنة) واختبارهم باستخدام الخطأ التربيعي الأصغري MAE لاستنتاج التركيبة المناسبة.

نستعرض فيما يلي إحصائيّات تبيّن دقّة محرّك التّوصية في التنبّؤ بتقييمات المستخدمين في النّظام، حيث تمّ حساب دقّة التقييمات بطريقة Cross validation وذلك بتقسيم مجموعة المعطيات Trip Advisor والمكوّنة من 22 ألف تقييم إلى مجموعتين: الأولى للتدريب تشكّل 80% من مجموعة المعطيات، فيما أجرينا الاختبار على الباقي من المعطيات (20%)، وبتكرار العمليّة 5 مرات وذلك بتغيير اختيار عيّنة التّدريب والاختبار من مجموعة المعطيات بشكلٍ عشوائيّ لتفادي مشاكل من نوع ارتباط التقييمات في مجموعة المعطيات المدروسة.

1.6.5 مقارنة النموذج المقترح مع النموذج التقليدي

بالاعتماد على الطريقة الموضحة سابقاً، قمنا بتدريب النموذج للحصول على الترتيب الأفضل للمعايير والذي تبين أنه:

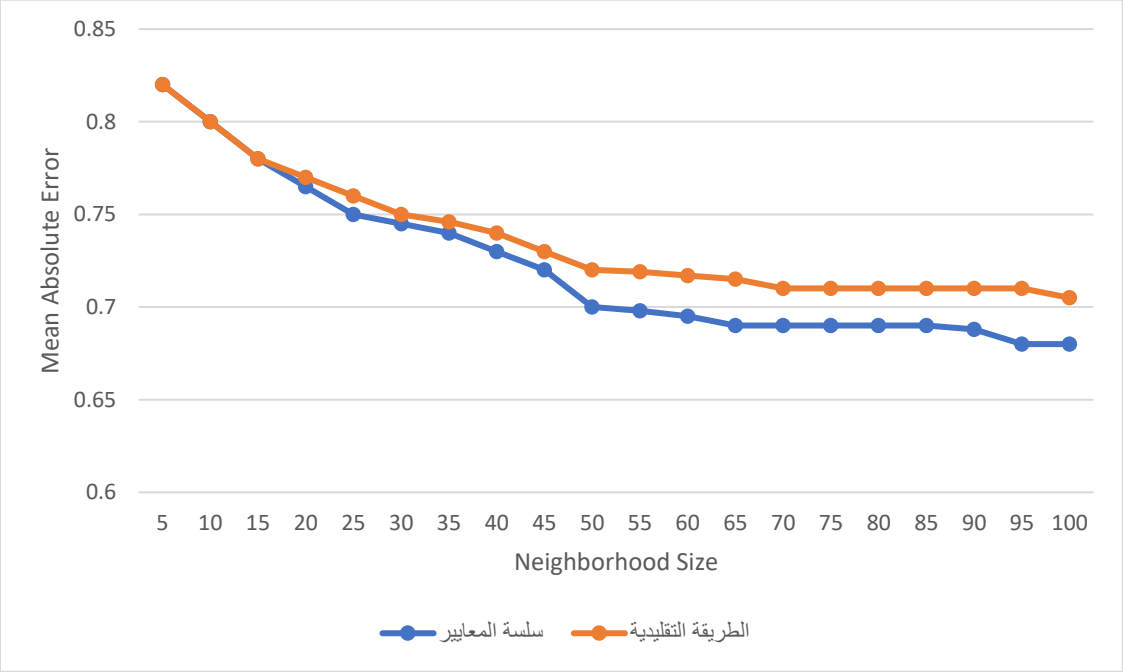
1. القيمة مقارنة بالمال المدفوع Value.
2. جودة خدمة الأعمال Business Service.
3. تجربة تسجيل الدخول للفندق Check in / Front Desk.
4. ملاءمة موقع الفندق Location.
5. جودة الغرف Rooms.
6. جودة الخدمة المقدمة Service.
7. نظافة الفندق Cleanliness.

بالمقارنة مع النموذج التقليدي الذي نعتمد فيه على الترشيح التعاوني بحجم عينة جوار هي 50 مستخدماً وباستخدام نفس مقياس الترابط Pearson، وجدنا أن قيمة الخطأ التربيعي الأصغري هي 0.7 (من 5 أي 14%) في سلسلة المعايير مقارنة مع 0.8 (من 5 أي 14.6%).

نبين فيما يلي تأثير اختيار حجم عينة الجوار على أداء نظام التوصية.

2.6.5 تأثير اختيار قيمة حجم عينة الجوار على أداء نظام التوصية

لدراسة تأثير اختيار قيمة الجوار على أداء النّظام قمنا بحساب دقّة النّظام في التنبّؤ بالتقييمات مع تغيير قيم عدد الجوار المعتمد، وذلك في حالتيّ سلسة المعايير والترّشيح التّعاونيّ اعتماداً على المستخدم، ويوضّح المخطّط التّالي النتائج الّتي حصلنا عليها:



الشكل 10 تناقص الخطأ في عمليّة التنبؤ مع ازدياد عدد الجوار.

نلاحظ انخفاض قيمة الخطأ في عمليّة التنبّؤ مع ازدياد عدد الجوار المعتمد في النّظام، حيث يكون هذا الانخفاض كبيراً عند القيم الصّغيرة لعدد الجوار، ومن ثمّ يتناقص تأثير زيادة عدد الجوار بعد عتبة معيّنة.

وعليه وبحسب النتائج السّابقة نجد أنّه يمكننا اعتماد قيمة فوق عتبة معيّنة بحيث نتمكّن من الحصول على دقّة مقبولة مع الحفاظ على أداء جيّد في عمليّة التّوصية من حيث تعقيد العمليّات الحسابيّة، وقد تمّ اعتماد عدد جوار 50 في النتائج اللّاحقة.

3.6.5 مقارنة الخوارزمية المقترحة مع خوارزمية Yong Z.

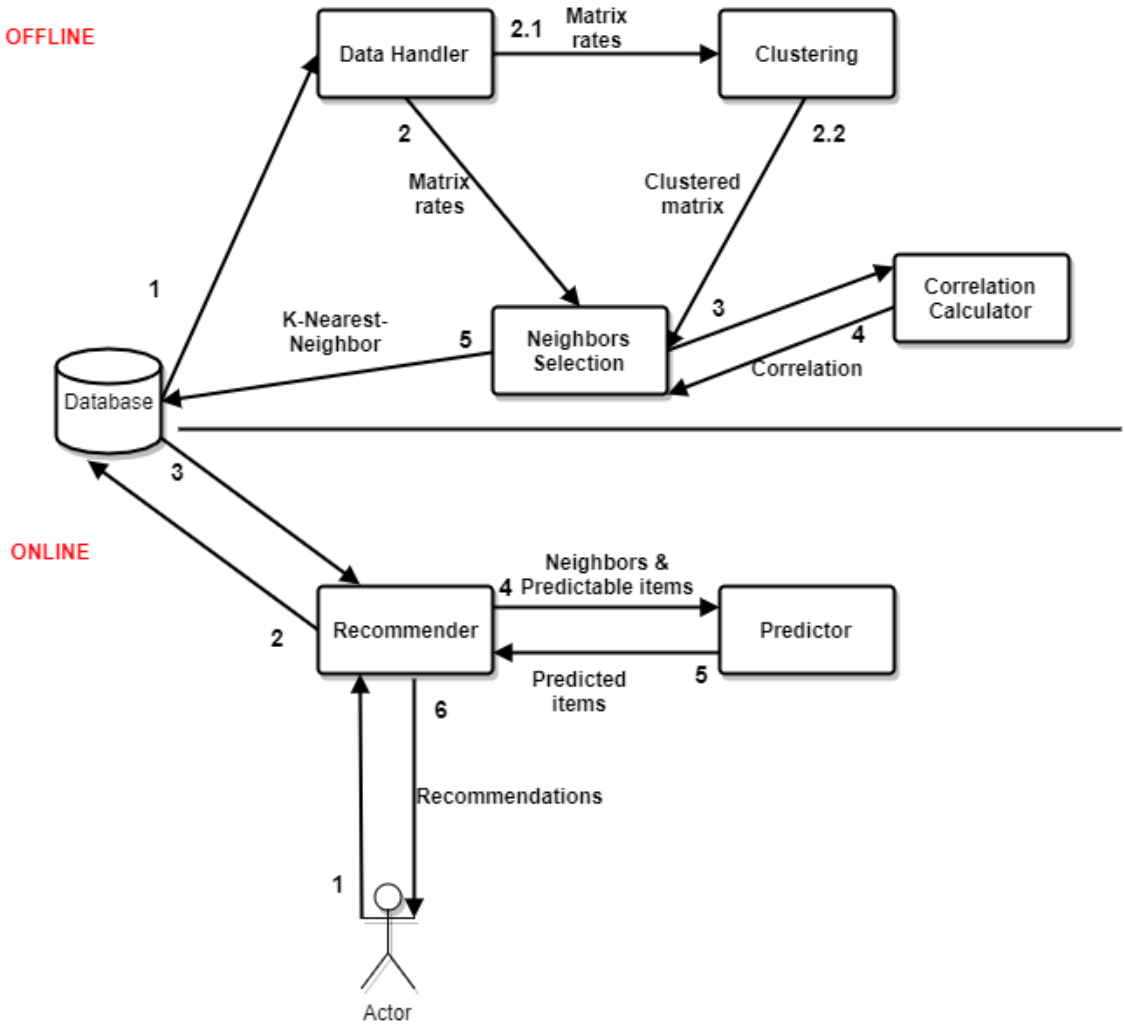
قام Yong Z. [60] بتطبيق خوارزميته على نفس المعطيات المدروسة في هذه الأطروحة وهي Trip Advisor Data Set. حققت خوارزميتنا المطروحة نتائجاً متقاربة بشكل كبيرة مع Yong Z. حيث حقق Yong Z. نتائج تقدر بخطأ تربيعي أصغري قيمته 0.78 (من أصل 5 أي 15.6%). نظراً لعدم تصريح Yong Z. بتفاصيل خوارزميته فإننا نتوقع الاختلاف الحاصل بسبب اختلاف طريقة حساب الترابط بين المعايير (حيث استخدمنا نحن مقياس Pearson للترابط)، وكما يمكن لاختلاف المقياس في حساب التشابه بين المستخدمين (حيث استخدمنا نحن مقياس Pearson المعدل).

7.5 المخطط الداخلي لعمل النظام البرمجي

إنّ إطار عمل النّظام يمكن تقسيمه ضمن شقيّن:

- **مرحلة تدريب**: وهي مرحلة تتمّ بمعزل عن عمل النّظام offline، يتمّ فيها إيجاد جوار كل غرض أو مستخدم ضمن النّظام باستخدام مقاييس التّرابط المختلفة، يتمّ تخزين خرج هذه المرحلة لتستخدم فيما بعد بإيجاد التّوصيات.
- **مرحلة تشكيل توصيات**: يتمّ فيها التنبّؤ بتقييمات المستخدمين للأغراض ضمن النّظام، وتوصية كل مستخدم بما يناسبه.

يوضّح المخطّط التّالي آليّة عمل النّظام، مبيّناً العمليّات في كلّ مرحلة من المراحل السّابقة:



الشكل 11 المخطط الداخلي لعمل النظام.

يوضّح المخطّط السّابق مراحل العمل في كلّ من المرحلتين المذكورتين سابقاً، مرحلة التدريب ومرحلة التّوصية، وفيما يلي نورد شرحاً للخطوات الخاصّة بكلٍّ مرحلةٍ كما هو مبيّن وفق المخطّط السّابق:

مرحلة التدريب: وتكون الخطوات ضمن هذه المرحلة كالتّالي:

1. تقوم طبقة التّعامل مع المعطيات بتشكيل مصفوفة (أو مكعب) تقييمات المستخدمين وفقاً للمعطيات المخزّنة في قاعدة معطيات النّظام.
2. يتمّ تمرير خرج المرحلة السّابقة وهي مصفوفة التقييمات إلى الوحدة المسؤولة عن انتقاء الجوار المناسب لكلّ مستخدم أو غرض، وكما هو مبيّنٌ في الشّكل السّابق فهناك خطوة اختياريّة بإجراء عمليّة العنقدة على المستخدمين أو الأغراض وذلك وفق الخطواتِ التّالية:
   1.2. تمرير مصفوفة التقييمات من طبقة التّعامل مع المعطيات إلى الطّبقة المسؤولة عن عمليّة العنقدة.
   2.2. تمرير مصفوفة التقييمات الجديدة بعد إجراء عمليّة العنقدة عليها إلى الطّبقة المسؤولة عن اختيار الجوار الخاص بكلّ عنقود.
3. إرسال المعلومات اللّازمة لحساب قيم التّرابط بين المستخدمين،، أو العناقيد إلى الوحدة المسؤولة عن هذه العمليّة.
4. تعيد هذه الوحدة قيم التّرابط المحسوبة وفق أحد المقاييس المنتقاة (وهو Pearson في حالتنا).
5. بعد أن تقومُ الوحدة المسؤولة عن اختيار أقرب k جار لكلّ مستخدم/عنقود، والّذين يحقّقون أعلى قيم ترابط مع المستخدم /العنقود الحاليّ، يتمّ تخزين هذا الجوار في قاعدة المعطيات ليستخدم فيما بعد بعمليّة إنشاء التّوصيات.

- **مرحلة تشكيل التوصيات**: وتكون الخطوات ضمن هذه المرحلة كالتّالي:
  1. يُطلب من النّظام تشكيل توصيات خاصّة بمستخدم معيّن.
  2. يتم تحصيل المعلومات اللّازمة والمتعلّقة بهذا المستخدم من قاعدة المعطيات، كالجوار الخاصّ به والأغراض القابلة للتنبّؤ.
  3. يتم تحصيل هذه المعلومات وتجهيزها ضمن بنى المعطيات المناسبة للمرحلة التالية.
  4. يتم إرسال هذه المعلومات المتعلّقة بجوار المستخدم، والأغراض القابلة للتنبّؤ إلى الوحدة المسؤولة عن حساب التنبّؤات.
  5. بعد أن يجري التنبّؤ بتقييم المستخدم للأغراض السّابقة وفق آليّات التنبّؤ المتّبعة في النّظام، تعيد هذه الوحدة التنبّؤات المحسوبة إلى الوحدة المسؤولة عن التّوصية.
  6. يجري توصية المستخدم بالأغراض الأكثر ملاءمة له، والتي تنبّأ لها النظام بأعلى التقييمات.

8.5 بيئة العمل البرمجية

تم تنجيز النّظام وقاعدة المعطيات ببيئة عمل .NET + SQL Server.

:.NET framework

تتميّزNET. بقوّتها في العمل وبوجود منصّة عمل IDE ممتازة (Visual Studio 2015)، حيث أنّ هذه المنصّة مدعومة من عملاق البرمجيّات Microsoft، حيث تتجهُ هذه الشّركة نحو تشغيل تطبيقات .NET. على نظام لينوكس عن طريق دعم منصّة Mono الّتي هي عبارة عن منصّة مفتوحة المصدر مبنيّة على بيئة عمل .NET. بحيث تسمح للمطورين ببناء تطبيقات غير مرتبطة بمنصّة التّشغيل cross-platform. وبالتّالي فإنّ المشكلة الأساسيّة المرتبطة ببيئة عمل .NET. والّتي تكمن بأنّها لا تعمل إلّا على windows قدْ تمّ تجاوزها، ممّا يجعل هذه البيئة مناسبة تماماً لتنجيز هذا النّظام خصوصاً بقدرتها الكبيرة على اختصار أجزاء كبيرة من العمل البرمجي ببعض الإجراءات والتّعليمات.

ويعدّ نظام إدارة قواعد البيانات SQL Server الخيار الأكثر توافقاً مع بيئة .NET. وهو الخيار المنصوح به عبر الموقع الرّسميّ ل Microsoft على الشّبكة العنكبوتيّة.

الفصل الخامس: الخاتمة والآفاق المستقبلية

1.6 الخاتمة

قمنا في هذه الأطروحة بدراسة أنظمة التوصية بنوعيها التقليدي وحيد المعيار ومتعددة المعايير، حيث انطلقنا من دراسة أنظمة دعم القرار متعددة المعايير لنصل إلى دراسة أنظمة التوصية كأنظمة دعم قرار للمستخدمين. قمنا بتطبيق إطار العمل المقترح من قبل Roy [16] ضمن سياق نظم التوصية متعددة المعايير. اقترح Roy منهجية وإطار عمل تقوم على أربعة أعمدة أساسية وهي:

- تحديد غرض القرار: وهو ما يقابل المادة المقترحة في نظم التوصية.
- عائلة المعايير: وهي اللبنة الأساسية في خوارزميتنا المقترحة والتي وصّفها Roy بوجوب تواجد ثلاثة شروط أساسية في وضعها وهي أن تكون هذه العائلة كاملة الترتيب وشاملة وغير زائدة عن الحاجة.
- نموذج التفضيل العام: حيث اعتمدنا على تابع تجميع خطي Linear Aggregation Model يربط بين تقييمات المستخدمين الإفرادية للانطلاق إلى التقييم الكلي.
- إجراء دعم القرار: وهو ما يربط بين الأعمدة الثلاثة السابقة للوصول إلى خوارزمية أو طريقة موحدة وواضحة لاتخاذ القرار متعدد المعايير.

قمنا باقتراح خوارزمية سميناها بـ "سَلسَلة المعايير Criteria Chaining" تقوم على فكرة ربط المعايير فيما بينها وتوقع التقييم الفردي باستخدام التقييمات الأخرى الفردية وتقييم المستخدم المتوقع على المعايير الأخرى التي تسبق التقييم الهدف ضمن الترتيب المدروس للمعايير.

تقوم الخوارزمية بدراسة كافة التراتيب الممكنة للمعايير (والتي تشكل $k!$ ترتيباً عند k معيار)، ثم الحصول على الترتيب الذي يعطي الخطأ الأصغري ضمن عينة الاختبار.

من الصعوبات التي واجهتنا خلال هذا البحث فقر الأبحاث ضمن مجال أنظمة التوصية متعددة المعايير وحداثة هذا النوع من الأنظمة ضمن الأدبيات العلمية، كما كان من الصعب تأمين مجموعة معطيات مفتوحة المصدر للعمل عليها في هذا البحث.

2.6 الآفاق المستقبلية

تضمن العمل ضمن الخوارزمية المقترحة بعض الطرق التقليدية المدروسة سابقاً مثل الترشيح التعاوني ومقياس Pearson لحساب الترابط.

يمكن العمل على الخوارزمية المقترحة ضمن طرق أخرى مثل الطرق التي تعتمد على المحتوى Content Based ومقاييس ترابط أخرى مثل Cosine Similarity وPearson المعدل الذي يأخذ بعين الاعتبار التناقضات بين المستخدمين، كما يمكن تطبيق بعض المقاييس التي طرحت في مشروع تخرج للطالب مجد العكة [63] تحت إشرافي المباشر والذي تم طرحه ضمن سياق هذه الأطروحة للمساعدة ببعض الأعمال فيها في المعهد العالي للعلوم التطبيقية والتكنولوجيا.

يمكن أيضاً العمل على تطوير الخوارزمية المقترحة لتأخذ تفضيلات كل مستخدم على حدة، أي تشكيل ترتيب معايير خاص بكل مستخدم بناءً على تفضيلاته وطريقة تقييمه للمواد السابقة.

تفتقر الأنظمة العالمية لنظم تقييم متعددة المعايير لصعوبة تجربة الاستخدام الناتجة عنها مما يؤدي إلى عزوف المستخدمين عن التقييم بشكل سهل. لتلافي هذه المشكلة، يمكن العمل على تحليل النصوص الموجودة ضمن المراجعات النصية واستخراج الميزات منها لتصبح هذه الميزات هي المعايير المدروسة في خوارزميتنا.

الملاحق

1.7 ملحق 1: أنظمة التوصية وحيدة المعيار

1.1.7 التّوصية باعتماد طريقة الجوار

تقوم أنظمة التّوصية المعتمدة على هذه الطريقة بأتمتة مبدأ ***نصيحة خبير***، حيثُ يعتمد الشّخص على رأي الأشخاص المتوافقين معه، أو المصادر الموثوقة لتقييم غرض ما (فلم، كتاب، مقال، أو غيرها...) ولتوضيح ذلك نستعرض المثال التّالي:

ليكن لدينا المستخدم بهاء والذي يريد أن يَحسِم قراره في مشاهدة فيلم "Titanic" من عدمه، يعلمُ بهاء أنّه وعمّار يملكان الذّوق نفسه في اختيار الأفلام، حيث أنّهما يتّفقان بالرّأي ببغضِهما لفيلم "The Matrix"، ويُفضّلان "Forrest- Gump"، لذا طلب نصيحته بشأن "Titanic". ومن جهة أخرى، وجدَ بهاء أنّه وصديقه خالد يملكان أذواقاً مختلفة فيما يتعلق بالأفلام، حيث يفضل خالد أفلام الإثارة وهو النوع الذي لا يفضله بهاء، لذا هو لا يهتمّ لرأي صديقه أو أنّه يأخذُ بعكسه. يبيّن الجدول التّالي تقييمات المستخدمين المذكورين أعلاهُ للأفلام المطروحة:

الجدول 2 مثال لتقييم أربعة مستخدمين لخمسة أفلام.

	The Matrix	Titanic	Die Hard	Forrest Gump	Wall-E
Majd	5	1		2	2
Ammar	1	5	2	5	5
Bahaa	2	?	3	5	4
Khaled	4	3	5	3	

2.1.7 مميزات اعتماد طريقة الجوار Neighborhood Approach

تمتازُ هذه الطّريقة بقدرتها على التقاط الارتباطات المحليّة في البيانات ضمن النّظام. بحيث يتمكّن نظام توصية أفلام يعتمد هذه الطّريقة من توصية فيلم يختلف عن الذّوق الاعتياديّ للمستخدم، وذلك في حال قامَ أحد جوار هذا المستخدم بتقييمه تقييماً جيّداً. قد لا تضمن هذه التّوصية نجاحاً أكيداً في ملائمة ذوق المستخدم، ولكنّها تُساعد في استكشاف أنماط جديدة للمستخدم توسّعُ نطاق تفضيلاتِه. يمكن تلخيص ميّزات هذه الطّريقة بما يلي: [62]

- **البساطة:** تعتبر هذه الطّريقة بديهيّة وسهلة التّنجيز، ففي أبسط أشكالها تحتاج فقط لمعايرة مُعامل عدد الجوار اللّازم للتنبّؤ بالتقييمات.
- **القدرة على التّبرير:** تقدّم هذه الطّريقة تبريراً حدسيّاً وواضحاً للتقييم المحسوب. فمثلاً، يمكن تقديم لائحة الأغراض المشابهة للغرض المتنبّأ بتقييمه مع تقييمات المستخدمين المختلفين لهذا الغرض كتبرير للمستخدم عن سبب التوصية

المطروحة، والّذي من شأنه أن يشكّل قاعدة لتشكيل نظام تفاعلي يتيحُ للمستخدم تحديد الجوار المناسب والّذي يمكن الاعتماد عليه في توصية أغراض جديدة.

- **الكفاءة:** يمتاز هذا النمط من الأنظمة بالكفاءة العالية، على العكس من الأنظمة المعتمدة على النّماذج model-based systems وهي الأنظمة التي تعتمد على بناء نموذج من المعطيات الموجودة بحيث يعتمد هذا النّموذج في توصية أغراض جديدة للمستخدمين، حيث أنّ طريقة الجوار لا تتطلّب مرحلة التدريب المكلّفة والواجب تكرارها خلال مجالات زمنيّة محدّدة. ولكن تكون فيها مرحلة التّوصية أكثر كلفة من تلك الأنظمة التي تعتمد على النّموذج. يتمّ في هذه الأنظمة إعادة حساب الجوار القريب لكل مستخدم في خطوة مستقلّة عن عمل النّظام، وعند الانتهاء من هذه الخطوة يتم اعتماد الجوار الجديد الّذي تمّ حسابه، ويتمّ تقديم التوصيات بشكل فوري دون إجراء حسابات الجوار أثناء عمليّة التّوصية، حيث يتمّ تخزين هذا الجوار والذي يحتاج لحجوم صغيرة من الذّاكرة، ممّا يجعل هذا النّمط فعّالاً في التّطبيقات الّتي تضمّ ملايين المستخدمين والأغراض.
- **الاستقرار:** تمتاز أنظمة التّوصية التي تعتمد طريقة الجوار باستقرارها وتأثّرها الطّفيف بالزّيادة المستمرّة للمستخدمين والأغراض والتقييمات، وهو الأمر الذي يمكن ملاحظته في التّطبيقات التّجاريّة الكبيرة. فعلى سبيل المثال، في الأنظمة التي تعتمد على الغرض في التّوصية، وحالما تمّ حساب التّشابه بين الأغراض، يمكن للنظام توصية أغراض للمستخدمين الجدد دون الحاجة لإعادة حساب التّشابه، وكذلك عندما يتمّ تقييم غرض جديد ضمن النّظام، نحتاج فقط لحساب التّشابه بين الغرض الجديد والأغراض الموجودة ضمن النّظام مسبقاً.

3.1.7 التنبّؤ بالتّقييم اعتماداً على المستخدم

تتنبّأ أنظمة توصية الجوار اعتماداً على المستخدم User-based neighborhood recommenders بالتقييم r_{ui} للمستخدم u من أجل غرض i باستخدام التقييمات المختلفة للمستخدمين المشابهين ل u والمتعلقة بالغرض i. لنفترض أنّنا نملك لكل مستخدم $v \neq u$ قيمة w_{uv} تمثل نسبة التّشابه بين المستخدمين u و v. ولتكن $\mathcal{N}(u)$ تمثّل k أقرب جار (مستخدم) للمستخدم u، وهم المستخدمين أصحاب أكبر نسب التّشابه مع المستخدم u، آخذين بعين الاعتبار أنّه فقط المستخدمين الذين قيّموا الغرض i هم الذين يدخلون في التنبّؤ بقيمة r_{ui}، وبالتالي تكون مجموعة k أقرب مستخدم ل u والذين قيّموا الغرض i هي المجموعة $\mathcal{N}_i(u)$. يمكن تقدير قيمة r_{ui} على أنّها متوسّط التقييمات التي حصل عليها الغرض i من مجموعة الجوار السّابقة [62]:

$$\hat{r}_{ui} = \frac{1}{|\mathcal{N}_i(u)|} \sum_{v \in \mathcal{N}_i(u)} r_{vi} \qquad \text{Eq29}$$

تكمن المشكلة في المعادلة السّابقة بأنّها لا تأخذ في عين الاعتبار اختلاف نسب التّشابه بين المستخدم الحالي مع الجوار المتعلّق به، فمثلاً بالعودة للمثال السّابق في الجدول 1 وباعتبار أنّ كلّ من عمار وخالد ينتمون إلى الجوار القريب من بهاء، سيكون من غير المنطقي اعتبار تقييم كلّ منهما بالتساوي للتنبّؤ بتقييم بهاء لفلم Titanic، حيث أنّ ذوق عمار أكثر تشابهاً مع بهاء من ذوق خالد. لذا من الممكن أن نضيف إلى كلّ جار يدخل في حساب التنبّؤ وزن خاص به وهو نسبة تشابه هذا المستخدم مع المستخدم الحالي وليكن w_{uv} هو وزن التّشابه بين المستخدمين u و v. وبالتالي سيكون شكل المعادلة [62]:

$$\hat{r}_{ui} = \frac{\sum_{v \in \mathcal{N}_i(u)} w_{uv}\, r_{vi}}{\sum_{v \in \mathcal{N}_i(u)} |w_{uv}|} \qquad \text{Eq30}$$

اقترحت مجموعة GroupLens نظام توصية بالتّرشيح التّعاونيّ باعتماد طريقة الجوار، حيث تقدّم GroupLens توصيات بالمقالات الإخباريّة للمستخدمين، وقد استخدمت في نظامها طريقة Pearson correlation لحساب أوزان التّشابه بين المستخدمين – سنستعرض هذ المقياس في الفقرة 4.3.2–، واعتمدت للتنبّؤ بتقييم مستخدم u لغرض i المعادلة التّالية [62]:

$$\hat{r}_{ui} = \overline{r_u} + \frac{\sum_{v \in \mathcal{N}_i(u)} w_{uv}\, (r_{vi} - \overline{r_v})}{\sum_{v \in \mathcal{N}_i(u)} |w_{uv}|} \qquad \text{Eq31}$$

حيث $\overline{r_v}$ هي متوسّط تقييمات المستخدم v، إن الإضافة الأهمّ في هذه العلاقة هي معالجتها مشكلة أنّ المستخدمين عادة ما يستفيدون من سلّم التقييمات بطرق مختلفة، فمثلاً في سلّم تقييم [1,5] قد يكون أعلى تقييم لمستخدم ما هو 3 وأقلّ تقييم 1، أمّا بالنّسبة لمستخدم آخر فقد يكون الأمر معاكس بحيث أنّه يعطي أفضل تقييم 5 والأسوء 3، وبذلك في حال أردنا الأخذ برأي المستخدم الأوّل بالنّسبة لغرض قيّمه بتقييم 3 سنجد أنّ هذا الغرض منصوحٌ به لأن تقييمه ينحرف عن متوسّط المستخدم بالاتّجاه الموجب، أمّا بالنّسبة للمستخدم الثّاني والّذي قيّم نفس الغرض بنفس التقييم فسنجد أنّ هذا الغرض غير منصوح به، وذلك لأنّ تقييمه ينحرف بالاتّجاه السّالب عن متوسّط المستخدم. ويكون الحلّ بأخذ متوسّط انحرافات الجوار عن متوسط تقييماتهم، ومن ثمّ إضافة هذا الانحراف إلى متوسّط تقييمات المستخدم الحالي للتنبّؤ بتقييمه للغرض المراد. تعبّر w_{uv} عن وزن التّرابط (التّشابه) بين المستخدمين u و v كما تم تشير إليها علاقة Pearson correlation: [64]

$$w_{uv} = \frac{\sum_{i=1}^{k} (r_{ui} - \overline{r_u})\, (r_{vi} - \overline{r_v})}{\sqrt{\sum_{i=1}^{k} (r_{ui} - \overline{r_u})^2 \sum_{i=1}^{k} (r_{vi} - \overline{r_v})^2}} \qquad \text{Eq32}$$

تمثّل k في المعادلة السّابقة عدد الأغراض المشتركة بين المستخدمين u, v، $\overline{r_u}$ هي متوسّط تقييمات المستخدم u، و r_{ui} هي تقييم المستخدم u للغرض i ويلاحظ من العلاقة السّابقة أنّ قيمة وزن التّرابط تنتمي للمجال $[-1, 1]$، سواءً التّرابط الإيجابي أو السّلبي.

4.1.7 حساب التّرابط

تمّ اعتماد العديد من المقاييس لحساب التّرابط بين مستخدمين أو غرضين، ونستعرض فيما يلي أبرز هذه المقاييس:

1. Cosine Similarity

مقياس لحساب التّرابط بين غرضين أو مستخدمين، حيث يتمّ حساب التّرابط بين مستخدمين u, v بتمثيل المستخدم u بشعاع x_u، والمستخدم v بشعاع x_v، وبالتّالي يمكن حساب التّرابط بين v و u وفق العلاقة التّالية [62]:

$$\cos(x_u, x_v) = \frac{x_u \, x_v}{\|x_u\| \, \|x_v\|}$$

$$CV(u,v) = \cos(x_u, x_v) = \frac{\sum_{i \in \mathrm{I}_{uv}} r_{ui} \, r_{vi}}{\sqrt{\sum_{i \in \mathrm{I}_u} {r_{ui}}^2 \;\; \sum_{j \in \mathrm{I}_v} {r_{vj}}^2}} \qquad \text{Eq33}$$

حيث تمثّل I_{uv} مجموعة الأغراض المشتركة بين المستخدمين u, v، بينما تمثّل I_u مجموعة الأغراض التي قام المستخدم u بتقييمها، وبنفس الطريقة يمكن حساب التّرابط بين غرضين.

يعاني هذا المقياس من مشكلة متوسّط وتباين التقييمات لكل من u, v.

2. Pearson Correlation

يعتبر من أشهر المقاييس المستخدمة في حساب التّرابط بين مستخدمين أو غرضين، حيث يمكّننا هذا المقياس من التخلّص من كثير من المشاكل، كحالة مستخدمان يقيّمان الأغراض ضمن النّظام وفق سلّمين مختلفين، فمثلاً ليكن لدينا مستخدم u ضمن نظام ذو سلّم تقييم [1,10]، ولكن u يستخدم المجال [4,7] فقط، بحيث يكون أفضل تقييم يعطيه هو 7 والأسوأ 4، وليكن المستخدم v يقيّم الأغراض ضمن المجال من [7,10]، بحيث أفضل تقييم لديه هو 10 والأسوأ هو 7، وبالتالي قد يتّفق هذان المستخدمان بالرأي على مجموعة الأغراض المشتركة بينهما، ولكنْ لن نكون قادرين على ملاحظة هذا الاتّفاق نتيجة اختلاف سلّم تقييم كلّ منهما، كحال المعهد العالي للعلوم التطبيقيّة والتكنولوجيا، مع باقي جامعات القطر، حيث أنّ تقييم 80% لطالب في المعهد العالي قد يكافئ 90% لنفس الطّالب في جامعة دمشق، ولكنْ من وجهة نظر أخرى فإن كلا المؤسستين تتفقان على أنّ هذا الطّالب هو طالب متفوّق. وبالتالي لا بدَّ من اكتشاف هكذا علاقات في المعطيات الموجودة ضمن النّظام.

$$w_{uv} = \frac{\sum_{i=1}^{k}(r_{ui} - \overline{r_u})\,(r_{vi} - \overline{r_v})}{\sqrt{\sum_{i=1}^{k}(r_{ui} - \overline{r_u})^2 \sum_{i=1}^{k}(r_{vi} - \overline{r_v})^2}} \qquad \text{Eq34}$$

تمثّل k في المعادلة السّابقة عدد الأغراض المشتركة بين المستخدمين u, v، ويُلاحَظ من العلاقة السّابقة أنّ قيمة وزن التّشابه تنتمي للمجال $[-1, 1]$، حيث أنّ الوزن السّابق يعبّر عن التّرابط بين المستخدمين، سواءً التّرابط الإيجابي أو السّلبي.

كما أنّ التّرابط السّالب بين مستخدمين يمكنْ الاستفادة منه في التنبُّؤ بالاتجاهِ العكسيّ فمثلاً لو كان لدينا مستخدمين بينهما ترابط عكسي قويّ، فيُمكننا الاعتماد على هذا التّرابط في التنبّؤ لأحدهما بتقييمات تعاكس الآخر.

يمكن استخدام نفس الفكرة في حساب التّرابط بين غرضين من النّظام i, j ولكن هذه المرة بمقارنة تقييمات المستخدمين اللذين قاموا بتقييم كلّ من هذين الغرضين وفق العلاقة التاليّة [64]:

$$w_{ij} = \frac{\sum_{u=1}^{k}(r_{ui} - \overline{r_i})\,(r_{uj} - \overline{r_j})}{\sqrt{\sum_{u=1}^{k}(r_{ui} - \overline{r_i})^2 \sum_{i=1}^{k}(r_{uj} - \overline{r_j})^2}} \qquad \text{Eq35}$$

تمثّل k في المعادلة السّابقة عدد المستخدمين الّذين قاموا بتقييم كل من الغرضين i, j، r_{ui} تقييم المستخدم u للغرض i، و$\overline{r_i}$ متوسّط تقييمات الغرض i.

3. Adjusted Cosine

تتمحور الفكرة الأساسيّة حول أنّه غالباً ما يكون التّباين بمقاييس تقييم الأشخاص أكبر من التّباين في مقاييس التقييمات الّتي تحصل عليها الأغراض، أيّ أنّ الأغراض الجيّدة غالياً ما تحصل على تقييمات عالية من أغلب المستخدمين، والأغراض السّيّئة تحصل على تقييمات منخفضة، وبالتّالي سيكون التّباين في التقييمات بالنّسبة لغرضٍ من هذهِ الأغراض ليسَ بالكبير مقارنةً بالتباين في تقييمات المستخدم، إذْ أنّ الطّيف الواسع من الأغراض المختلفة في جودتِها ينتج عنه تباين في تقييم المستخدم الواحد لهذه الأغراض. وعليه أثناء حساب تشابه الأغراض سيكون من الأجدر مقارنة التقييمات بالنسبة لمتوسّط تقييمات المستخدمين، بدلاً من اعتبارها بالنّسبة لتقييمات الأغراض وهو الحال في المقياس Pearson. يعتبر هذا المقياس تعديلاً للمقياس Pearson Item Similarity.

$$w_{ij} = \frac{\sum_{u=1}^{k}(r_{ui} - \overline{r_u})\,(r_{uj} - \overline{r_u})}{\sqrt{\sum_{u=1}^{k}(r_{ui} - \overline{r_u})^2 \sum_{i=1}^{k}(r_{uj} - \overline{r_u})^2}} \qquad \text{Eq36}$$

تُمثّل k في المعادلة السّابقة عدد المستخدمين الّذين قاموا بتقييم كلٍّ من الغرضين i,j، r_{ui} تقييم المستخدم u للغرض i، و$\overline{r_u}$ متوسّط تقييمات المستخدم u.

4. Frequency-Weighted Pearson Correlation

تكمنُ الفكرة الأساسيّة وراءَ هذا المقياس في أنَّ تقييمات مستخدمين اثنين على الأغراض المتّفق عليها من قبل الجمهور، غالباً ما تكون غير مفيدة في تحديد التّرابط بين هذين المستخدمين، فمثلاً يتّفق النّاس عامةً على محبّة فلم Titanic، وبالتّالي فلو أنَّ مستخدمين اثنين قاما بتقييم هذا الفلم تقييماً عالياً، فلا يمكننا استنتاج أنَّ هذين المستخدمين متشابهان، إذ أذ أنّ محبّة هذا الفلم هي أمر متّفق عليه، بينما لو اتّفق مستخدمان على تقييم Saw تقييماً عالياً، فيمكن من خلاله استنتاج أنّ هذين المستخدمين محبّان لأفلام الرّعب وبالتّالي هناك ترابط في مكان ما بين هذين المستخدمين. وبالمثل، فإنّ المستخدم الّذي دائماً ما يقيّم الأغراض بالاتجاه نفسه، أيّ دائماً ما يعطي تقييمات عاليّة جداً للأغراض، أو بالعكس، سيكون أقلّ فائدة تنبّؤيّة من ذلك المستخدم الّذي غالباً ما تختلف تقييماته من غرض لآخر.

يُسمّى المنهج الّذي يتناول هذه القضيّة Inverse User Frequency، وهو مشتقٌّ من مفهوم Inverse Document Frequency (IDF)، والمستخدم في مجال استرجاع المعلومات، حيث يتم إضافة وزن λ_i لكل غرض i، وهو لوغاريتم نسبة عدد المستخدمين في النّظام إلى عددِ المستخدمين الّذين قاموا بتقييم الغرض i.

$$\lambda_i = \log\frac{|U|}{|U_i|}$$

وعند حساب التّرابط بين مستخدمين تستخدم العلاقة التّالية [62]:

$$w_{uv} = \frac{\sum_{i=1}^{k} \lambda_i \, (r_{ui} - \overline{r_u}) \, (r_{vi} - \overline{r_v})}{\sqrt{\sum_{i=1}^{k} \lambda_i \, (r_{ui} - \overline{r_u})^2 \sum_{i=1}^{k} \lambda_i \, (r_{vi} - \overline{r_v})^2}} \qquad \text{Eq37}$$

وبالطّريقة نفسها يمكن استخدام العلاقة السّابقة لحساب التّرابط بين غرضين من النّظام.

5.1.7 مصداقية أوزان التّرابط

عادةً ما يتمّ حساب أوزان التّرابط بناءً على عددٍ قليل من التقييمات المشتركة للأغراض بين المستخدمين، بسبب ضآلة عدد التقييمات الّتي يقوم بها المستخدمون للأغراض الموجودة ضمن النّظام مقارنةً بعددِ المستخدمين وعدد الأغراض.

وبحسب المثال الوارد في [62]: فليكنْ لدينا نظام مكوّن من 10,000 تقييم من قبل 1,000 مستخدم ل 100 غرض، وبافتراض توزّع منتظم لهذه التقييمات، عندها ستكون حسابات التّشابه بين المستخدمين قائمة على تقييم لغرضٍ مشتركٍ واحدٍ بشكلٍ وسطيّ. فإذا كانت هذه التقييمات المشتركة القليلة بين مستخدمين متساوية بالقيمة، عندها سيعتبر هذا المستخدمان متشابهين بشكلٍ كاملٍ. وبالتّالي إذا كانت أذواق المستخدمين متناقضة حقيقةً، سيؤدّي ذلك إلى توصيات ضعيفة غير مناسبة.

تتمحور الحلول المقترحة لتجاوز المشكلة السّابقة حول تخفيض وزن نسبة التّشابه بين مستخدمين عندما تكون هذه النسبة محسوبة بناءً على عددٍ قليلٍ من التقييمات المشتركة. وكان من الحلول المقترحة وفقاً ل [62] تصغير هذا الوزن بنسبة تتعلّق بعدد الأغراض المشتركة بين المستخدمين في حال كان هذا العدد أصغر من عتبة معيّنة $\gamma > 0$:

$$\acute{w}_{uv} = \frac{\min\{|I_{uv}|, \gamma\}}{\gamma} \times w_{uv} \qquad \text{Eq38}$$

بناءً على المعادلة السّابقة فإنّه يتمّ ضرب وزن التّشابه بنسبة عدد الأفلام المشتركة إلى عتبة محدّدة، وفي حال تجاوز عدد الأفلام المشتركة في حساب هذا الوزن العتبة المحددة، عندها يتمّ احتساب الوزن الأصلي دون تصغيره، وبحسب [62] فإنّ اختيار $\gamma \geq$ 25 كفيلٌ بتحسين دقّة نظام التّوصية، وأنّ القيمة 50 أعطتْ أفضل نتيجة، وبشكلٍ عام فإنّ قيمة هذا المعامل ترتبط بالبيانات المدروسة ويمكن تحديدها بشكلٍ مستقلّ من نظامٍ لآخر.

6.1.7 اختيار الجوار

يؤثّر اختيار عدد الجوار الخاص بكلّ مستخدم أو كلّ غرض (التّرشيح التّعاونيّ باعتماد المستخدم أو باعتماد الغرض)، بالإضافة إلى طريقة اختيار هذا الجوار على جودة محرّك التّوصيّة، حيث يتمّ اختيار الجوار الّذي يستخدم في توصية الأغراض عادة من خلال مرحلتين:

1- ترشيح مسبق للجوار يبقي فقط على المرشّحين الأكثر قُرباً.

2- مرحلة التنبّؤ والّتي يتمّ فيها اختيار أفضل المرشّحين لهذا التنبّؤ.

1.6.1.7 الترشيح المسبق للجوار

عادةً ما يستحيل تخزين كل أوزان التّشابه بين كلّ زوج من المستخدمين أو الأغراض، وذلكَ نظراً لمحدوديّة التّخزين وخاصّةً في أنظمة التّوصية الكبيرة والّتي تضمّ ملايين المستخدمين والأغراض، كما أنّ تخزين كلّ هذه القيمة يعتبرُ هدر غير مبرّر لوحدات التّخزين، حيث أنّه يتمّ استخدام الجوار الأكثر قرباً فقط عند التنبّؤ بالتقييمات. وعليه فمن الضّروريّ إجراء ترشيح مسبق للجوار لتصبح طريقة اعتماد الجوار قابلة للتطبيق من خلال تقليل عدد أوزان التشابه المراد تخزينها، وتحديد عدد الجوار المرشّح استخدامه في التنبّؤ بالتقييمات. ويمكن تنفيذ هذا التّرشيح المسبق بعدة طرق: [62]

- **ترشيح أعلى n جار**: يتم الاحتفاظ بلائحة تضمّ n أقرب جار لكلّ مستخدم أو غرض مع الأوزان الموافقة لكلّ منهم. ولتفادي المشاكل المتعلّقة بالفعاليّة والدقّة يجب اختيار قيمة n بدقّة، فباختيار n كبيرة جداً سيتطلّب ذلك حجم كبير من الذّاكرة لتخزين لوائح الجوار، كما أنّ عمليّة التنبّؤ بالتقييمات ستكون بطيئة. بالمقابل فإنّ اختيار قيم صغيرة جداً لـ n قدْ يؤدّي إلى نقص في التّغطية الخاصّة بنظام التّوصية والّذي سيسبّب بألّا يتمّ توصية بعض الأغراض أبداً.
- **التّرشيح بالتعتيب**: تعتمد هذه الطريقة على الاحتفاظ بكافة الجوار الّذين تتجاوز أوزان تشابههم عتبة محدّدة w_{min}، وذلك بدلاً من الاحتفاظ بعددٍ ثابتٍ من الجوار لكلّ مستخدم أو غرض، وبينما تعتبر هذه الطريقة أكثر من سابقتها إلّا أنّ اختيار قيمة مناسبة للعتبة يعتبر من المسائل الصّعبة.
- **ترشيح القيم السّالبة**: عموماً تكون أوزان التّرابط الموجبة أكثر وثوقيّة من تلك السّالبة. فحدسيّاً، يعود ذلك إلى أنّ أوزان التّرابط الموجبة القويّة بين مستخدمين تكون مؤشّر جيّد على انتماء هذين المستخدمين إلى مجموعة مشتركة (مراهقين، محبّي الخيال العلمي...)، وبينما تدلّ القيم السّالبة على انتماء المستخدمين إلى مجموعتين مختلفتين، فإنّها لا تعطي أي معلومة عن مدى اختلاف المجموعتين، أو عن كون هاتين المجموعتين تتناسبان مع صنف معيّن من الأغراض. وقد بيّن الدراسات التجريبيّة أنّ اعتبار القيم السّالبة للارتباطات لم يعطي أي تحسين في الدقّة.

2.6.1.7 الجوار في التنبّؤات

بعد الحصول على لائحة الجوار الخاصّة بكل مستخدم أو غرض، يتمّ التنبّؤ بتقييم جديد اعتماداً على أقرب k جار، وهم أصحاب أعلى أوزان تشابه مع المستخدم أو الغرض الحالي، وتبقى المسألة الأهمّ هنا وهي تحديد قيمة مناسبة لـ k.

بيّنت الدراسات والتّجارب [62] أنّ المنحني البيانيّ لدقّة التنبّؤات بالنسبة لقيم k المختارة له شكل مقعّر، بمعنى آخر عند اختيار قيم صغيرة لـ k مثلا $k < 20$، غالباً ما تكون دقّة التنبّؤات ضعيفة، ومع زيادة قيم k، يزداد عدد الجوار المساهم في التنبّؤات وبالتّالي يُمكن التّخلّص من التّباين الخاصّ بكلّ مستخدم وزيادة دقّة التنبّؤات. أخيراً فقدْ تبيّن أنّ زيادة عدد الجوار المستخدم في التنبّؤ يؤثّر سلباً على دقّة النّظام مثلاً $k > 50$، وذلك نظراً لحقيقة أنّ العلاقات القويّة لبعض الجوار تفقد قوّتها نظراً لإدخال الكثير من الجوار ذو الارتباط الضّعيف مع المستخدم أو العنصر، ويبقى أنّ تحديد القيمة المثلى لـ k يكون بالتّجريب cross-validation، واختيار القيمة الأمثل وفق البيانات الّتي يتعامل معها النّظام.

7.1.7 عنقدة المستخدمين والأغراض

يصعب عادة اكتشاف التّشابه بين المستخدمين والأغراض، كوننا لا نملك الكثير من المعلومات عن الأزواج مستخدم–غرض وذلك نظراً لقلّة التقييمات من قبل المستخدمين وتبعثرها، إذْ يمكن أن نجد غرضين ينتميان للنوع ذاته (فلمين من النمط الكوميدي) دون أن يكون هناك أيّ مستخدم قامَ بشراء هذين الغرضين أو تقييمهما معاً. كما يمكن أن يكون لمستخدمَين الذّوق نفسه دون أن يشتركا بأيّ من الأغراض الّتي قاموا بشرائها أو تقييمها.

اقترح [65] طريقة للتعامل مع هذه الإشكاليّة (القيم الفارغة في مصفوفة المستخدمين والأغراض)، وهي عنقدة الأغراض أو المستخدمين أو كليهما معاً، وذلك باستخدام مقاييس التّشابه مثل Cosine Similarity.

فليكن لدينا المثال التّالي والّذي يستعرض تقييمات بعض المستخدمين لمجموعة من الأفلام:

الجدول 3 تقييم مجموعة من المستخدمين لمجموعة من الأفلام.

	HP1	HP2	HP3	TW	SW1	SW2	SW3
A	4			5	1		
B	5	5	4				
C				2	4	5	
D		2					3

بعد إجراء عمليّة عنقدة للأفلام الواردة في المثال السابق سنحصل على الجدول التالي:

الجدول 4 نتيجة عنقدة مجموعة أفلام HP ضمن عنقود واحد، وكذلك مجموعة SW.

	HP	TW	SW
A	4	5	1
B	4.67		
C		2	4.5
D	2		3

بعد إجراء عمليّة العنقدة على الأغراض، تصبح أعمدة المصفوفة في الجدول 3 تمثّل عناقيد من الأغراض بينما تمثّل الأسطر المستخدمين، حيث تصبح قيمة المدخل للمستخدم u والعنقود C هي متوسّط تقييمات u للأغراض الّتي قام باستعراضها وتنتمي للعنقود C. وفي حال لم يملك المستخدم أي تقييمات لأغراض عنقود C فسيبقى المدخل المقابل فارغاً.

يمكن بعد إتمام عمليّة العنقدة السّابقة على الأغراض إجراء عمليّة عنقدة للمستخدمين فتصبح المصفوفة السّابقة هي تقابل عناقيد من المستخدمين مع عناقيد من الأغراض. كما يمكن تكرار العمليّة السّابقة أي عنقدة عناقيد المستخدمين والأغراض مراراً وتكراراً حتّى نحصل على عدد مقبول من العناقيد، أي تصغير أبعاد المسألة للحدّ المقبول.

نستعرض فيما يلي شرحاً لخوارزميّة العنقدة K-Means، وهي خوارزميّة لتجميع البيانات استناداً إلى خصائصها وسماتها، وتتمّ عمليّة التّجميع من خلال تقليلِ المسافات بين البيانات ومراكز التّجميع (cluster centroid).

تتلخّص آليّة عمل هذه الخوارزميّة بالخطوات التالية:

1- حساب إحداثيّات مراكز التّجميع: عند التّكرار الأوّل يتمُّ اختيار المراكز بشكلٍ عشوائيٍّ (كاعتماد نقاط عشوائيّة من المعطيات على أنّها مراكز ابتدائيّة للعناقيد)، وفي التّكرارات اللّاحقة يتمّ حساب المراكز على أنّها متوسّط النّقاط المنتمية إلى العُنقود.

2- حساب المسافة بين كلّ البيانات ومراكز التّجميع.

3- تجميع البيانات وتنظيمها في مجموعاتٍ بناءً على أقلِّ المسافات بين المركز ونقاط البيانات.

4- إعادة تنفيذ الخطوات من $1 - 3$ حتى الوصول إلى حالةِ الثّبات أو الوصول إلى عددٍ معيّن من التكرارات نقف بعده.

يعتمدُ أداء هذه الخوارزميّة على المواقع الأوّليّة لمراكز التّجمّع (Centroid)، ومن المستحسن تنفيذ هذه الخوارزميّة عدّة مرّات مع اختلاف المراكز في كلّ مرّة عن المرّات السّابقة.

تفاصيل تنجيز الخوارزميّة:

- حساب مراكز التّجميع (مراكز العناقيد): هو متوسّط النّقاط المنتمية إلى هذا العنقود.
- حساب المسافة: تمّ اعتماد المسافة الإقليديّة في تنجيز الخوارزميّة، وهناك خيارات أخرى يمكن اعتمادها مثل مسافة هامينغ وغيرها.
- ستقوم الخوارزميّة بالتّجميع اعتماداً على المقاييس المذكورة سابقاً.
- عادةً معظم التجميع يتمّ في المراحل الأولى.
- تعقيد الخوارزميّة هو $O(n.k.t.m)$ حيث:
  - n هي عدد السّجلّات.
  - k هي عدد العناقيد المراد تشكيلها.
  - t هي عدد مرّات التّكرار حتى الحصول على العناقيد النهائيّة.
  - m هي عدد واصفات السّجل الواحد (تتعلّق بالتّابع المعتمد في حساب المسافة).

ملاحظة: يمكن تطبيق هذه الخوارزميّة عندما يمكن تعريف تابع مسافة بين أيّ سجلّين، وتعريف متوسّط عنقود من السّجلات.

بعد إجراء عمليّة العنقدة السّابقة وللحصول على تقييم مستخدم u لغرض i: [65]

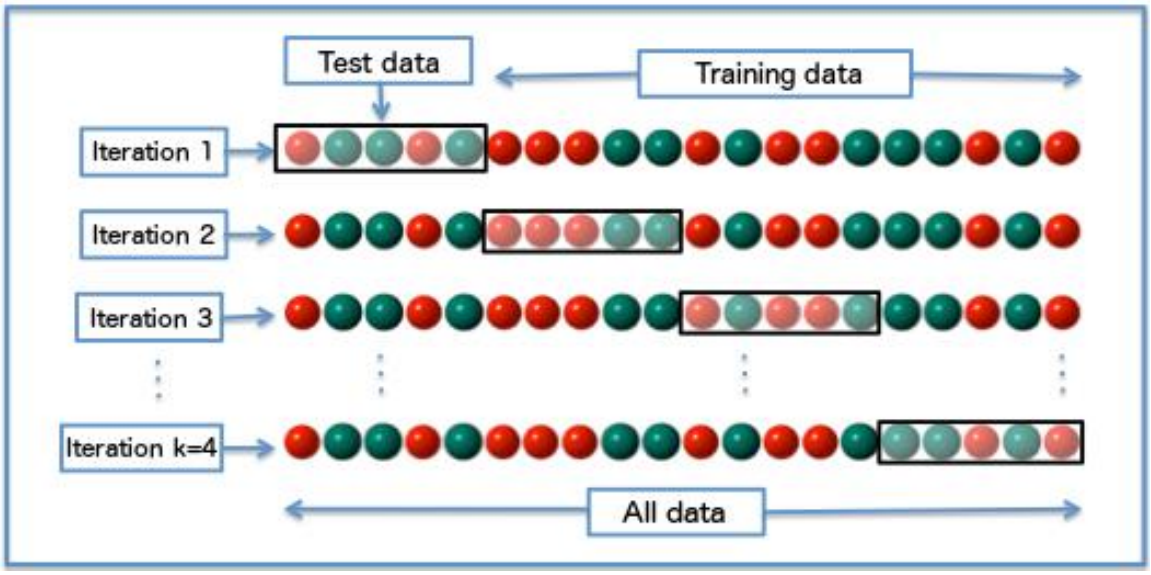
- نوجد العنقود الّذي ينتمي له كل من u و i، وليكن العنقودان C و D على الترتيب.
- إذا كان المدخل في مصفوفة العناقيد المقابل ل C و D غير فارغ، عندها تستخدم هذه القيمة لتقدير تقييم u للغرض i في المصفوفة الأساسيّة قبل العنقدة.
- إذا كان هذا المدخل فارغاً، عندها نستخدمُ إحدى طرق تقدير التقييمات التي تحدّثنا عنها، كالتّرشيح التّعاونيّ اعتماداً على المستخدم وغيرها.

إنّ أبرز المشكلات الّتي تعاني منها عمليّة العنقدة هي انخفاض دقّة التقييمات نتيجة اعتماد متوسّط تقييمات المستخدم للأغراض في العنقود الواحِد، وإسناد هذه القيمة كتقييم هذا المستخدم لباقي أغراض العنقود والّتي قد يكون عددها كبيراً، وبالتّالي وفي حال قيّم المستخدم غرض واحد ضمن العنقود فسيتمّ إسناد التقييم ذاته لباقي أغراض هذا العنقود وهو الأمر الذي يؤثّرُ سلباً على دقّة التقييمات.

8.1.7 اختبار النّظام Testing

إنَّ الهدف من عمليّة الاختبار هو قياس مدى جودةِ النظام وقدرتِه على التنبّؤ بتقييمات المستخدمين للأغراض المختلفة، ولعلّ التقنية الأشهر في هذا المجال هي التحقق التّصالبيّ Cross-Validation [66]:

وهي تقنيّة معياريّة في الاختبارات ضمن مجال التّنقيب في المعطيات Data Mining، تساعدُ المطوّر في اختبار النّماذج المبنيّة، حيث تقومُ هذه التقنيّة على مبدأ تقسيم مجموعة المعطيات المتوافرة للدراسة إلى قسمين، عادةً ما يكون (70%-90% للتدريب وبناء النموذج، فيما يتم استخدام (10% - 30%) المتبقيّة لاختبار النموذج المبنيّ. ثمّ يتمّ تكرار العمليّة السّابقة في تقسيم المعطيات إلى قسمين عدداً من المرات بحيثُ يتمّ تغيير عيّنة التدريب والاختبار في كلّ مرةٍ وذلك تفادياً لمشاكل من نوع ارتباط المعطيات، كأن تكون عيّنة الاختبار المنتقاة مرتبطة بعيّنة التدريب وبالتّالي نحصل على نتائج غير دقيقة للنّموذج المبني، وبالتّالي يتمّ تكرار العمليّة عدداً من المرات واعتماد متوسّط النتائج الإحصائيّة، ويوضّح الشّكل التّالي آليّة لتقسيم المعطيات.



الشكل 12 مرحل تقسيم المعطيات في تقنية Cross-Validation.

كما هو موضّح في الشّكل السّابق فقدْ جرى اختيار كلٍّ من العيّنتين بشكلٍ تكراريٍّ مروراً على مجموعة المعطيات بحيث لا تتقاطع أيّ من المجموعتين مع الأخرى، وتجدر الإشارة هنا إلى أنّ هذه الطّريقة ليست الوحيدة في تقسيم مجموعة المعطيات، إذْ يمكن تقسيمها بشكلٍ عشوائيّ بحيث يتمّ اختيار نسبة معيّنة منها عشوائيّاً للتدريب، واعتماد ما تبقّى للاختبار في كلِّ تكرار.

لاختبار أداء نظام التّوصية هناك عاملان أساسيّان يجب مراعاتهما، وهما دقّة التنبّؤات النّاتجة عن النّظام، وقدرة النّظام على تغطية التقييمات الّتي تُطلب منه، وسنبيّن كلّ منهما فيما يلي:

1. التّغطية Coverage

هي مقياس لنسبة الأغراض الّتي تنبّأ نظام التّوصية بتقييماتها. وتُعتبر نسبة الأغراض الّتي تمكّن النّظام من التنبّؤ بتقييماتها، إلى الأغراض المطلوب التنبّؤ بتقييماتها هي المقياس الأساسيّ للتّغطية.

من العوامل الّتي تؤدّي إلى انخفاض هذه النّسبة هو اعتماد عددٍ قليلٍ من الجوار للمستخدمين والأغراض ضمن النّظام. ويمكن أنْ نعزي عدم مقدرة النّظام على التنبّؤ ببعضِ التقييمات إلى أنّ الأغراض المطلوب التنبّؤ بتقييماتها قدْ لا تملكْ أي تقييمات سابقة ضمن النّظام، أو أنّ عدد قليل من المستخدمين والّذين لا ينتمون لجوار المستخدم النّشط هم من قاموا بتقييمها فقط.

2. الدّقّة Accuracy

تمّ اقتراح عدّة مقاييس لتقييم دقّة نظام التّوصية من النّاحية الإحصائيّة، والّتي تقيّم دقّة هذا النّظام من خلال مقارنة القيم العدديّة للتنبّؤات مع التقييمات الفعليّة للمستخدمين وذلك بالنسبة للأغراض الّتي نملك عنها المعلومتين السّابقتين. ويعتبر المقياس Mean Absolute Error (MAE) من أشهر المقاييس في هذا المجال والّذي يعتمد على حساب متوسّط الفروقات بين التنبّؤات والقيم الحقيقيّة بالقيمة المطلقة.

المراجع

[1] F. Ricci ،"Travel Recommender Systems "،*IEEE Intelligent Systems* ،6 رقم ،17 المجلد ،pp. 55-57 .2002 ،

[2] Gavalas, Kenteris, "A web-based pervasive recommendation system for mobile tourist guides," *Personal & Ubiquitous Computing*, vol. 15, no. 7, pp. 759-770, 2011.

[3] M. J. a. D. B. Pazzani ،Content-based recommendation systems ،Berlin: Springer ،2007 ،pp. 325-341.

[4] J. B. e. a. Schafer ،Collaborative filtering recommender systems, The adaptive web ،Berlin: Springer ،2007 ، pp. 291-324.

[5] N. H. Carlos A. Gomez-Uribe, "The Netflix Recommender System: Algorithms, Business Value, and Innovation," *ACM Transactions on Management Information Systems (TMIS)*, vol. 6, no. 4, 2016.

[6] G. Adomavicius and A. Tuzhilin ،"Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions "،*IEEE Transactions on Knowledge and Data Engineering* ،6 رقم ،17 المجلد ، pp. 734-749 .2005 ،

[7] J. A. Konstan ،"Introduction to recommender systems: Algorithms and evaluation "،*ACM Transactions on Information Systems (TOIS)* ،1 رقم ،22 المجلد ،pp. 1-4 .2004 ،

[8] G. A. a. Y. Kwon ،"New recommendation techniques for multicriteria rating systems "،*IEEE Intelligent Systems* ،3 رقم ،22 المجلد ،pp. 48-55 .2007 ،

[9] N. M. a. C. Costopoulou ،"Analysis and classification of multi-criteria recommender systems "،*World Wide Web: Internet and Web Information Systems* ،4 رقم ،10 المجلد ،pp. 415-441 .2007 ،

[10] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl ،"Grouplens: An open architecture for collaborative filtering of netnews تأليف "،*The 1994 ACM conference on Computer supported cooperative work* ، .1994

[11] Gediminas Adomavicius, Alexander Tuzhilin, "Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734-749, 2005.

[12] "Tf-idf weighting," Stanford, [Online]. Available: https://nlp.stanford.edu/IR-book/html/htmledition/tf-idf-weighting-1.html. [Accessed 3 8 2017].

[13] Zitnick, C.L., Kanade, T., "Maximum entropy for collaborative filtering.," in *AUAI '04: Proc. of the 20th Conf. on Uncertainty in Artificial Intelligence*, Virginia, 2004.

[14] Kantor, Francesco Ricci · Lior Rokach · Bracha Shapira · Paul B., "Content-based Recommender Systems: State of the Art and Trends," in *Recommender Systems HandBook*, Springer US, 2010, pp. 73-107.

[15] PAZZANI, DANIEL BILLSUS. MICHAEL J., "User Modeling and User-Adapted Interaction," in *User modeling for adaptive news access.*, Kluwer Academic, 2000, pp. 147-180.

[16] B. R. a. M. R. McCord ،Multicriteria methodology for decision aiding ،Springer .1996 ،

[17] J. Figueira, S. Greco, and M. Ehrgott ،"Multiple criteria decision analysis: state of the art "،*Springer Verlag* ،
.2005

[18] R. L. Keeney and H. Raiffa ،"Decisions with multiple objectives: Preferences and value tradeoffs،" Cambridge Univ Pr.1976 ،

[19] E. Triantaphyllou ،"Multi-criteria decision making methods: a comparative study "،*Kluwer Academic Pub* ،
.2000

[20] E. Jacquet-Lagrze and Y. Siskos ،"Preference disaggregation: 20 years of mcda experience "،*European Journal of Operational Research* ،2 رقم ،130 المجلد ،pp. 233-245 .2001 ،

[21] Y. S. a. C. Z. P. M. Pardalos ،Advances in multicriteria analysis ،Springer .1995 ،

[22] K. Deb ،"Multi-objective optimization." Search methodologies ،Boston: Springer ،2014 ،pp. 403-449.

[23] B. Roy and D. Bouyssou ،Aide multicrit`ere `a la d´ecision: m´ehodes et cas ،Paris: Economica Paris .1993 ،

[24] R. Burke ،"Knowledge-based recommender systems "،*Encyclopedia of Library and Information Systems* المجلد ،
4 رقم ،12pp. 331-370 .2000 ،

[25] Adomavicius, G & ,.Kwon, Y ،"Multi-Criteria Recommender Systems تأليف "،*Recommender Systems Handbook* ،Springer ،2015 ،pp. 847-880.

[26] Amatriain, X., Basilico, J. ،"Netflix recommendations: Beyond the 5 stars،" 2012 .[متصل] .Available: https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429 .
.[تاريخ الوصول 2018].

[27] Siskos, Yannis, Evangelos Grigoroudis, and Nikolaos F. Matsatsinis ،""UTA methods." Multiple criteria decision analysis: State of the art surveys "،*Springer* ،pp. 297-334 .2005 ،

[28] Breese, J., Heckerman, D., Kadie, C. ،"Empirical analysis of predictive algorithms for collaborative filtering "،
تأليف*In Proc. of the 14th Conference on Uncertainty in Artificial Intelligence* ،San Francisco, CA .1998 ،

[29] Manouselis, N., Costopoulou, C. ،"Experimental analysis of design choices in multiattribute utility collaborative filtering "،*International Journal of Pattern Recognition and Artificial Intelligence* ، المجلد 21، رقم 2، pp. 311-332 .2007 ،

[30] Tang, T., McCalla, G. ،"The pedagogical value of papers: a collaborative-filtering based paper recommender "، *Journal of Digital Information* .2009 ، المجلد 10، رقم 2،

[31] Maneeroj, S., Samatthiyadikun, P., Chalermpornpong, W., Panthuwadeethorn, S., Takasu, A ،"Ranked criteria profile for multi-criteria rating recommender "،*Information Systems, Technology and Management* ،pp. 40-51 ، .2012

[32] Sarwar, B., Karypis, G., Konstan, J., Riedl, J. ،"Item-based collaborative filtering recommendation algorithms تأليف "،*In: Proc. of the 10th International Conference on World Wide Web* ،New York, NY, USA ، .2001

[33] Boulkrinat, S., Hadjali, A., Mokhtari, A. ،"Towards recommender systems based on a fuzzy preference aggregation تأليف "،*In: 8th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-13)* ،Atlantis .2013 ،

[34] Palanivel, K., Siavkumar, R. ،"Fuzzy multicriteria decision-making approach for collaborative recommender systems "،*International Journal of Computer Theory and Engineering* ،1 رقم ،2 المجلد ،pp. 57-63 .2010 ،

[35] Nilashi, M., Ibrahim, O., Ithnin, N. ،"Hybrid recommendation approaches for multi-criteria collaborative filtering "،*Expert Systems with Applications* ،8 رقم ،41 المجلد ،pp. 3879-3900 .2014 ،

[36] Sahoo, N., Krishnan, R., Duncan, G., Callan, J ،" Research note-the halo effect in multicomponent ratings and its implications for recommender systems: The case of yahoo! movies "،*Information Systems Research* ، المجلد ، 1، رقم 23pp. 231-246 .2012 ،

[37] Si, L., Jin, R. ،"Flexible mixture model for collaborative filtering تأليف "،*In: Proc. of the 20th International Conference on Machine Learning* .2003 ،

[38] M. Ehrgott ،Multicriteria optimization ،Springer Verlag .2005 ،

[39] Lakiotaki, K., Tsafarakis, S., Matsatsinis, N. ،"UTA-Rec: a recommender system based on multiple criteria analysis تأليف "،*In: Proc. of the 2008 ACM conference on Recommender systems* ،ACM New York, NY, USA .2008 ،

[40] Siskos, Y., Grigoroudis, E., Matsatsinis, N ،UTA methods ،Springer .2005 ،

[41] S .& .F. C. b. S. PhD ،"KENDALL'S TAU .[متصل] "،Available: http://stanfordphd.com/KendallsTau.html [تاريخ الوصول 9 2018].

[42] Akhtarzada, A., Calude, C., Hosking, J. ،"A multi-criteria metric algorithm for recommender systems "، *Fundamenta Informaticae* ،1 رقم ،110 المجلد ،pp. 1-11 .2011 ،

[43] Chow, C., Liu, C. ،"Approximating discrete probability distributions with dependence trees "،*IEEE Transactions on Information Theory* ،3 رقم ،14 المجلد ،pp. 462-467 .1968 ،

[44] Lee, H., Teng, W. ،"Incorporating multi-criteria ratings in recommendation systems تأليف "،*In: IEEE International Conference on Information Reuse and Integration* .2007 ،

[45] Lee, W., Liu, C., Lu, C. ،"Intelligent agent-based systems for personalized recommendations in internet commerce "،*Expert Systems with Applications* ،4 رقم ،22 المجلد ،pp. 275-284 .2002 ،

[46] J. Schafer ،"Dynamiclens: A dynamic user-interface for a meta-recommendation system تأليف "،*In: Proc. of the Workshop on the next stage of recommender systems research at the ACM Intelligent User Interfaces Conf* ، .2005

[47] Adomavicius, G., Sankaranarayanan, R., Sen, S., Tuzhilin, A. ،" Incorporating contextual information in recommender systems using a multidimensional approach "،*ACM Transactions on Information Systems (TOIS)* ،1 رقم ،23 المجلد ،pp. 103-145 .2005 ،

[48] Jannach, Dietmar & Zanker, Markus & Fuchs, Matthias ،"Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations "،*Information Technology & Tourism* ،14 المجلد ،pp. 119-149 .2014 ،

[49] Jameson, A., Smyth, B. ،"Recommendation to groups "،*Lecture Notes in Computer Science 4321* .2007 ،

[50] Lops, P., De Gemmis, M., Semeraro, G. ،" Content-based recommender systems: State of the art and trends "،. *Recommender systems handbook* ،pp. 73-105 .2011 ،

[51] Lousame, F., Sánchez, E. ،" Multicriteria predictors using aggregation functions based on item views تأليف "،. *Intelligent Systems Design and Applications (ISDA)* .2010 ،

[52] Lakiotaki, K., Matsatsinis, N., Tsoukias, A. ،"Multicriteria user modeling in recommender systems "،*IEEE Intelligent Systems* ،2 رقم ،26 المجلد ، pp. 64-76 .2011 ،

[53] Liu, L., Mehandjiev, N., Xu, D.L. ،"Multi-criteria service recommendation based on user criteria preferences "، تأليف*In: Proceedings of the fifth ACM conference on Recommender systems* ،ACM .2011 ،

[54] Jannach, D., Karakaya, Z., Gedikli, F. ،"Accuracy improvements for multi-criteria recommender systems "، تأليف*In: Proceedings of the 13th ACM Conference on Electronic Commerce* ،ACM .2012 ،

[55] Freund, Y., Iyer, R., Schapire, R., Singer, Y. ،" An efficient boosting algorithm for combining preferences "، *The Journal of Machine Learning Research* ،4 المجلد ، pp. 933-969 .2003 ،

[56] Nguyen, H., Haddawy, P. ،"DIVA: applying decision theory to collaborative filtering تأليف "،*In: Proc. of the AAAI Workshop on Recommender Systems* ،Madison, WI .1998 ،

[57] Saaty, T. ،" Optimization in integers and related extremal problems "،*McGraw-Hill* .1970 ،

[58] Schmitt, C., Dengler, D., Bauer, M. ،"Multivariate preference models and decision making with the maut machine تأليف "،*In: Proc. of the 9th International Conference on User Modeling (UM 2003)* .2003 ،

[59] D. Brandt ،" How service marketers can identify value-enhancing service elements "،*Journal of Services Marketing* .1988 ،

[60] Y. Zheng ،"Criteria Chains: A Novel Multi-Criteria Recommendation Approach تأليف "،*Proceedings of the 22nd International Conference on Intelligent User Interfaces* ، ACM New York, NY, USA .2017 ،

[61] Jesse Read, Bernhard Pfahringer, Geoff Holmes, and Eibe Frank ،"Classifier chains for multi-label classification "،*Machine learning* ،3 رقم ،85 المجلد ،pp. 333-359 .2011 ،

[62] Kantor, Francesco Ricci · Lior Rokach · Bracha Shapira . Paul B., "A Comprehensive Survey of Neighborhood-based Recommendation Methods," in *Recommender Systems Handbook*, Springer US, 2010, pp. 107-145.

[63] مجد العكة، ابراهيم الابراهيم، مصطفى دقاق، "بناء نظام توصية رقمي،" المعهد العالي للعلوم التطبيقية والتكنولوجيا، دمشق، 2017.

[64] Jonathan L. Herlocker, Joseph A. Konstan, Al Borchers, and Hohn Riedl, "An Algorithmic Framework for performing collaborative filtering," in *international ACM SIGIR conference on Research and development in information retrieval*, California, 1999.

[65] Jure Leskovec, Anand Rajaraman, Jeff Ullman, "Recommendation Systems," in *Mining of Massive Datasets*, Cambridge University Press, 2012, pp. 307-341.

[66] "Cross-Validation (Analysis Services - Data Mining)," Microsoft, 1 3 2017. [Online]. Available: https://docs.microsoft.com/en-us/sql/analysis-services/data-mining/cross-validation-analysis-services-data-mining. [Accessed 25 8 2017].

[67] Kantor, Francesco Ricci · Lior Rokach · Bracha Shapira · Paul B., "Advances in Collaborative Filtering," in *Recommender Systems Handbook*, Springer US, 2010, pp. 145-187.

[68] Yehuda Koren, Robert Bell, Chris Volinsky, "Matrix Factorization Techniques for Recommender Systems," *Computer*, vol. 42, no. 8, 2009.

[69] "http://www.mono-project.com/," Mono Project, 2017. [Online]. Available: http://www.mono-project.com/docs/about-mono/.

[70] "ASP.NET Data Access Options," Microsoft, [Online]. Available: https://msdn.microsoft.com/en-us/library/ms178359(v=vs.110).aspx.

[71] Mikeli,A.,Apostolou, D.,Despotis,D. ،"A multi criteria recommendation method for interval scaled ratings "، تأليف*Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM International Joint Conferences* .2013 ،

[72] Jiang, T., Tuzhilin, A ،"Segmenting customers from population to individuals: Does 1-to-1 keep your customers forever "،؟*IEEE Trans. on Knowl. and Data Eng* ،p. 1297–1311 .2006 ،

[73] Aciar, S., Zhang, D., Simoff, S., Debenham, J. ،" Informed recommender: Basing recommendations on consumer product reviews "،.*IEEE Intelligent systems 22* .2007 ،*(3)*

[74] Jameson, A., Smyth, B. ،"Recommendation to groups "،*Lecture Notes in Computer Science 4321* .2007 ،