

Syrian Arab Republic

Ministry of Higher Education

Syrian Virtual University



الجمهورية العربية السورية

وزارة التعليم العالي

الجامعة الافتراضية السورية

تصميم وتنفيذ نظام توصية على موقع ويب للأفلام

## Design and Implementation a Recommendation System on Movies Website

بحث مقدم لنبيل درجة الماجستير

في تقانات الويب

إشراف الدكتورة

سيرا أستور

إعداد الطالبة

لينا محمد عماد موصللي

Lina\_114886

2024

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

رَبِّ أَوْزِعْنِي أَنْ أَشْكُرَ نِعْمَتَكَ الَّتِي أَنْعَمْتَ عَلَيَّ وَعَلَى وَالَّذِي وَأَنْ أَعْمَلَ صَالِحًا  
تَرْضَاهُ وَأَدْخِلْنِي بِرَحْمَتِكَ فِي عِبَادِكَ الصَّالِحِينَ

الحمد لله الذي أعاذه في كل لحظة من اللحظات التي مرت بي خلال حياتي  
عموماً وخلال مسيرة العمل على هذه الرسالة بشكل خاص

الحمد لله الذي وفقني لتعلم ما تعلمته خلال السنين الماضية ووفقني للوصول  
إلى هذه المرحلة

الحمد لله الذي جعلني أستشعر عنایته بي كما شعر بها سيدنا موسى عليه السلام  
عندما قال له " قَالَ لَا تَخَافَا إِنِّي مَعَكُمَا أَسْمَعُ وَأَرَى "

الحمد لله الذي جعلني أتجاوز كل ما مر بي خلال مراحل العمل على هذه  
الرسالة وجب خاطري في إتمامها

اللهم علمنا ما ينفعنا وانفعنا ما عملتنا وزدنا علماً

# كلمة شكر

مهما تقدمنا وفتحت أمامنا الطرق ووصلنا لكل ما نحلم به، علينا أن نتذكر من  
كانوا سبب بنجاحنا، من ساندنا وأمسك بيدها للاستمرار، من وجودهم حفزنا  
وشجعنا، فمهما عبرنا لهم فالكلمات قليلة

لهذا أتقدم بالشكر والتقدير إلى الكادر التعليمي في الجامعة الافتراضية وإلى  
أساتذتنا الكرام في اختصاص تقانات الويب وأخص بالذكر الأساتذة الأفضل  
الذين تكروا بتحكيم دراستي، وتقديم العون الصادق دون تأخير

الأستاذ الدكتور محمد الجنيدى والأستاذ الدكتور محمد مازن مصطفى..

تسابق الكلمات وتتزاحم العبارات لتنظم عقد الشكر الذي لا يستحقه إلا أنت ..  
إليك يا من كان لها قدم السبق في ركب العلم والتعليم .. إليك يا من بذلت ولم تنتظري  
العطاء .. إليك يا من كنت بمثابة الأم بالنسبة لي .. إليك يا من رافقني عناء هذا  
الإنجاز بكل صبر .. إليك أهدي عبارات الشكر والتقدير

الدكتورة القديرة

سيراً أستور

# اللهم إلهي

إلى أول من أحبني.. إلى من انتظرت مجئي لهذه الدنيا بكل ما تملكه من حب.. إلى من تجرعت كأس الشقاء مرأً لتسقيني رحيق السعادة.. إلى التي صحت بالكثير من أجل أن أحيا.. إلى الشمس التي أنارت دربي.. إلى من أرى من خلال ثغرها الباسم جمال الكون ولذته.. إلى الصدر الذي يضماني كلما ضاقت بي الدنيا، وأحاطت بي المخاطر.. إلى قمرى الذي لا يغيب وإلى من روحها تعانق روحها الأبدى.. إلى أعز وأغلى مخلوق عندي

إلى أمي

إلى أعظم رجل في حياتي.. إلى الرجل الذي جعلني فتاة مدللة.. إلى من حمل قلبي على أكف السعادة، ورسم الضحكة على شفتي، وسقاني من حنانه وفيض عطائه، حتى وقفت على قدمي.. إلى من علمني معنى الحياة.. وأمسك بيدي على دروبها.. إلى قوتي وعزمي وسندى في هذه الدنيا.. إلى القدوة التي لطالما اقتديت بها.. إليك يا أحن وأطيب وأروع أب في هذا

الوجود

إلى أبي

إلى من قاسموني اللحظات بحلوها ومرها.. إلى من لا يمكنني العيش بدونهم..

إلى عيناي.. إخوتي ثائر و محمد

إلى من كانت كتفي الثابت في أصعب الأوقات.. إلى من لم تمل لحظة من الاستماع لي وال الوقوف إلى جانبي.. إلى حبيبة قلبي وأختي

إلى صديقة دربي.. رهام عاجي

إلى التي لطالما دعمتني وصدقت أنني أستطيع النجاح دوماً.. إلى التي لطالما كانت الركن المهدئ في حياتي.. إلى التي جمعتني بها الصدفة لتصبح الأحب إلى قلبي..

إلى صديقة دربي.. لين عمران

إلى الذي كان أول داعم لي في مسيرتي المهنية.. إلى الذي كان بمثابة الأب في توجيهاته وفي دعمه لي.. إلى الشخص الذي لم أكن ما أنا عليه الآن لولا ثقته وتشجيعه..

إلى الأستاذ المهندس عامر سياف

إلى الذي كان الداعم لي في عملي.. إلى الذي تولى الإشراف على تقديمي في هذه الرسالة.. إلى الذي لم يمل ولم يشك للحظة أن ب�能وري إنجاز هذا العمل.. إلى الشخص الذي لم يتوازن عن المبادرة بأية نصيحة.. والذي لم أستطع لولاه إتمام هذا العمل..

إلى الزميل المهندس مجد الزبيبي

إلى من ساروا معي في دروب العلم والحياة.. إلى من قضيت معهم أحلى الأيام..

إلى مياس زوجة أخي وأصدقاءي

إلى إخوتي وعائلتي وأصدقائي

إلى كل من وقف إلى جنبي

إليكم جميعاً أهدي هذا العمل المتواضع

## فهرس المحتويات

2 .....	الإهداء
6 .....	فهرس المحتويات
8 .....	فهرس الأشكال
9 .....	ملخص الرسالة
10 .....	المقدمة
11 .....	تاريخ أنظمة التوصية
12 .....	لماذا يجب علينا نستخدم أنظمة التوصية؟
13 .....	أمثلة على موقع الأفلام التي تحتوي على أنظمة توصية
15 .....	مشكلة البحث
15 .....	الهدف
17 .....	1- الفصل الأول: الدراسة النظرية
17 .....	1-1- دراسة حول موقع الأفلام ومتطلباتها
17 .....	1-1-1- وجهة نظر المستخدم
18 .....	1-1-2- من وجهة نظر المبرمج
19 .....	1-2- تعريف نظام التوصية
21 .....	1-3- أنظمة التوصية
21 .....	1-3-1- أنواع أنظمة التوصية
35 .....	1-3-2- مقارنة بين أنظمة التوصية: نحو تحسين تجربة المستخدم في البيئة الرقمية
39 .....	1-3-3- استخدام أكثر من نظام توصية
40 .....	1-3-4- سلبيات أنظمة التوصية
42 .....	2- الفصل الثاني: الدراسات المرجعية والأعمال المتعلقة
42 .....	2-1- مراجعة الدراسات السابقة
43 .....	2-2- الأعمال المتعلقة
43 .....	43 ..... مشكلة التوصية Netflix Prize
45 .....	45 ..... YouTube

3- الفصل الثالث: القسم العملي .....	48
3-1 التحليل .....	48
3-1-1 الهدف من المشروع .....	48
3-1-2 متطلبات المشروع .....	49
3-2 التصميم .....	53
3-2-1 مخطط حالة الاستخدام Use Case Diagram .....	53
3-2-2 ERD Diagram .....	58
3-2-3 مخطط الصف Class Diagram .....	62
3-3 التجيز .....	63
3-3-1 الأدوات البرمجية المستخدمة .....	63
3-3-2 مراحل العمل على نظام التوصية .....	69
3- الفصل الرابع: النتائج .....	88
4-1 واجهات الويب .....	88
4-1-1 واجهات الويب للزائر .....	89
4-1-2 واجهات الويب للتسجيل وتسجيل الدخول .....	93
4-1-3 واجهات الويب للمستخدم .....	94
4-1-4 واجهات المدير .....	95
5- الفصل الخامس: الآفاق المستقبلية .....	99
الخلاصة Conclusion .....	99
المراجع .....	100

## فهرس الأشكال

الشكل 1: انتشار وسائل الترفيه ..... 11
الشكل 2: خوارزميات التوصية ..... 20
الشكل 3: المصفوفة R ..... 21
الشكل 4: تحسين الترتيب عن الحالة الأولية ..... 45
الشكل 5: انتشار منصات التواصل الاجتماعي ..... 47
الشكل 6: أشهر لغات البرمجة 2023 حسب IEEE Spectrum ..... 66
الشكل 7: بنية Django ..... 67
الشكل 8: مراحل إعداد البيانات والتدريب الأولى ..... 70
الشكل 9: مصفوفة التشابه لعينة من البيانات ..... 75
الشكل 10: المخطط التدفقى لنهج التوصية الهجين ..... 77
الشكل 11: مراحل نموذج التوصية الأول ..... 78
الشكل 12: متوسط الأصوات إلى عدد الأصوات ..... 79
الشكل 13: مراحل نموذج التوصية الثاني (القسم الأول) ..... 81
الشكل 14: مراحل نموذج التوصية الثاني (القسم الثاني) ..... 81
الشكل 15: مصفوفة التشابه لـ The Gentlemen ..... 83
الشكل 16: المخطط التدفقى لإظهار البيانات للمستخدمين ..... 87

## فهرس الجداول

الجدول 1: مقارنة بين أنظمة التوصية المختلفة ..... 38
الجدول 2: أمثلة تفوق النموذج الهجين على النموذج القائم على المحتوى ..... 85
الجدول 3: أمثلة تفوق النموذج القائم على المحتوى على النموذج الهجين ..... 86

## ملخص الرسالة

يعاني مشاهدو الأفلام من تحديات في إيجاد محتوى يلبي تطلعاتهم على منصات العرض، مما يؤثر على جاذبية التجربة. يهدف هذا المشروع إلى تحسين هذه التجربة من خلال تطوير نظام توصية فعال. يتضمن النظام نموذجين، أحدهما يعتمد على نظام تصفية قائم على المحتوى والآخر يدمج بين التصفية التعاونية والتصفية القائمة على المحتوى. بحيث نهدف من هذا المشروع إلى بناء نظام توصية قادر على توفير توصيات فعالة وملائمة وذات صلة بما يبحث عنه المستخدم. ومن ثم بناء موقع ويب لعرض معلومات الأفلام والممثلين والتوصيات المتخصصة للمستخدمين بواجهات بسيطة سهلة التعامل مما يعزز تجربة المستخدم ويلبي متطلباته.

سنبدأ في مقدمة البحث بتحديد المشكلة التي يدور حوله بحثنا، والهدف الذي نتطلع لتحقيقه في نهاية البحث. كما يتضمن على لمحه تاريخية حول أنظمة التوصية وبداية ظهورها وأمثلة من مختلف مجالات تطبيقها ولاسيما في مجال موضع الأفلام عالمياً.

وستنطرق في الفصل الأول لدراسة حول موضع الأفلام ومتطلباتها وذلك من خلال تحديد المعايير التي تميز موقع أفلام محدد عن الآخر. ومن ثم سنقوم بالتوسيع في شرح أنظمة التوصية، أنواعها بما في ذلك ميزات كل نظام توصية وسلبياته، المقارنة بين أنظمة التوصية ونختم هذا الفصل بالطرق إلى سلبيات أنظمة التوصية بشكل عام.

يليه سنتناول في الفصل الثاني مجموعة من بعض الدراسات المرجعية التي ساهمت في تعزيز معارفنا حول أنظمة التوصية وأنواعها وخوارزمياتها. بالإضافة إلى الأعمال المتعلقة التي كانت تجارب ملهمة لنا في بحثنا هذا.

فيما بعد سنقوم في الفصل الثالث بتفصيل الإجراءات التي قمنا بها في مراحل بناء نظام التوصية والتقنيات والمكاتب المستخدمة في القسم العملي وما يتربى على ذلك من اختبار للنتائج.

ومن ثم سنعرض واجهات الويب التي سيعامل معها كل من مستخدمي الموقع في الفصل الرابع وأخيراً سنذكر بعضاً من الآفاق المستقبلية والمراجع التي استندنا عليها في بحثنا.

## المقدمة

إن وسائل الترفيه أصبحت رائجة في عصرنا هذا ومن أبسط هذه الوسائل وأكثرها رواجاً هي الأفلام.

الأفلام محبوبة من الجميع كباراً وصغاراً، ذكوراً وإناثاً، تجمعهم تحت مظلتها باختلاف أعرافهم وألوانهم ومواطねهم. نحن جميعاً متصلين بطريقة ما ببعضنا البعض عبر هذه الوسيلة المذهلة. ومع ذلك، فإن الأمر الأكثر إثارة للاهتمام هو حقيقة أن خياراتنا ومجموعاتنا فريدة من حيث تفضيلات الأفلام. يحب بعض الأشخاص الأفلام حسب نمطها سواء كانت تتنمي للإثارة أو الرومانسية أو الخيال علمي، بينما يركز البعض الآخر على الكادر سواء كانوا ممثلين أو مخرجين. عندما نأخذ كل هذه العوامل في الحسبان، فإننا نجد من الصعب جدًا تعليم فيلم والجسم بأن الجميع سيرغب بمشاهدته. ولكن مع كل ما تم طرحه، لا يزال يُرى أن الأفلام المماثلة تحظى بإعجاب مجموعة معينة من الأشخاص.

وفي ظل الانتشار الواسع للإنترنت على الأجهزة الذكية من حواسيب وهواتف النقالة، أصبحت منصات مشاهدة الأفلام والمسلسلات عبر الانترنت تغزو العالم وتتنافس فيما بينها من حيث الدقة، سرعة توفير المحتوى، أخذ حقوق النشر والحصرية وغيرها الكثير من الميزات.

لذلك أصبح من اللازم إيجاد أنظمة قوية لنشر محتوى الترفيه من مسلسلات وأفلام لشدة الطلب عليها. ومن أبرز الأمثلة عند التحدث عن أنظمة التوصية، YouTube وNetflix .. الخ



الشكل 1: انتشار وسائل الترفيه

## تاريخ أنظمة التوصية

أنظمة التوصيات لها تاريخ طويل، حيث تعود جذورها إلى أوائل التسعينيات. تتضمن بعض المعالم الرئيسية في تطوير أنظمة التوصية ما يلي (1):

- أنظمة التصفية التعاونية الأولى: استندت أنظمة التوصية ببدايةً إلى تقنيات التصفية التعاونية، والتي تحل سلوك المستخدم لتقييم توصيات مخصصة. كان نظام GroupLens، الذي تم تطويره في جامعة مينيسوتا في عام 1992، من أوائل أنظمة التصفية التعاونية الناجحة. ومن الأنظمة الأولى الأخرى كان Tapestry، الذي تم تطويره في Xerox PARC في عام 1993. وRingo، الذي تم تطويره في MIT في عام 1996.
- أمازون والتصفية التعاونية من عنصر إلى عنصر: في أواخر التسعينيات، أطلقت أمازون نظام التوصيات الخاص بها، والذي كان يعتمد على التصفية التعاونية من عنصر إلى عنصر. ركز هذا النهج على أوجه التشابه بين العناصر الفردية بدلاً من أوجه التشابه بين المستخدمين، وأصبح نهجاً شائعاً في صناعة التجارة الإلكترونية.

- جائزة Netflix: في عام 2006، أطلقت Netflix مسابقة بجائزة قيمتها مليون دولار لتحسين دقة نظام التوصيات الخاص بها بنسبة 10%. حفزت هذه المنافسة البحث والتطوير الكبير في مجال أنظمة التوصيات. استخدم الفريق الفائز مجموعة من التقنيات، بما في ذلك عامل المصفوفة وطرق التجميع.
- التوصيات الاجتماعية والتوصية القائمة على المحتوى: في أول عقد من القرن الواحد والعشرين، أصبحت التوصيات الاجتماعية شائعة بشكل متزايد، مع العديد من أنظمة التوصية التي تتضمن بيانات الشبكات الاجتماعية لتقديم توصيات أكثر تخصيصاً. أصبحت التصفيّة القائمة على المحتوى، والتي تحلّ خصائص العناصر لتقديم التوصيات، أكثر شيوعاً أيضاً.
- التعلم العميق والأنظمة الهجينة: في السنوات الأخيرة، تم تطبيق تقنيات التعلم العميق مثل الشبكات العصبية على أنظمة التوصية، مما أدى إلى تحسينات كبيرة في الدقة. أصبحت الأنظمة الهجينة، التي تجمع بين تقنيات التوصية المتعددة، أكثر شيوعاً أيضاً.

## لماذا يجب علينا نستخدم أنظمة التوصية؟

هناك العديد من الأسباب التي تجعل استخدام أنظمة التوصية أمراً أساسياً في أي موقع. ومنها:

1. التخصيص: يمكن أن تقدم أنظمة التوصية توصيات مخصصة مصممة لتناسب تفضيلات وسلوك المستخدم الفردي. يمكن أن يؤدي ذلك إلى تحسين تجربة المستخدم وزيادة التفاعل مع المنتج أو الخدمة.
2. زيادة المبيعات والإيرادات: من خلال التوصية بالمنتجات أو المحتوى ذي الصلة للمستخدمين، يمكن لأنظمة التوصية زيادة المبيعات والإيرادات للشركات. على سبيل المثال، ذكرت أمازون أن نظام التوصيات الخاص بها كان مسؤولاً عما يصل إلى 35% من مبيعاتها في عام 2017.
3. فعالة من حيث التكلفة: يمكن أن تكون أنظمة التوصية فعالة من حيث التكلفة للشركات، حيث يمكنها المساعدة في تقليل الحاجة إلى حملات التسويق المكلفة من خلال استهداف المستخدمين بتوصيات مخصصة.

4. تحسين رضا العملاء: من خلال تقديم توصيات مخصصة، يمكن لأنظمة التوصية تحسين رضا العملاء ولائهم. من المرجح المستخدمون إلى منتج أو خدمة إذا شعروا أنه تتم تلبية احتياجاتهم وتفضيلاتهم.

5. تقليل الحمل الزائد للمعلومات: في عالم اليوم، غالباً ما يكون هناك الكثير من المعلومات التي يمكن للمستخدمين معالجتها بمفردهم. يمكن أن تساعد أنظمة التوصية في تصفية المعلومات وتحديد أولوياتها، مما يسهل على المستخدمين العثور على ما يبحثون عنه.

6. الاكتشاف المحسن: يمكن أن تساعد أنظمة التوصية المستخدمين على اكتشاف منتجات وخدمات ومحفوظات جديدة ربما لم يجدوه بطريقة أخرى. يمكن أن يؤدي ذلك إلى تجربة مستخدم أكثر تنوعاً وإثراءً.

7. قابلية التوسيع: يمكن تحجيم أنظمة التوصية بسهولة للتعامل مع كميات كبيرة من البيانات ونشاط المستخدم، مما يجعلها مناسبة للاستخدام في منصات التجارة الإلكترونية الكبيرة وموقع التواصل الاجتماعي وغيرها من الخدمات عبر الانترنت.

بشكل عام، تعد أنظمة التوصية أداة قوية لتحسين تجربة المستخدم وزيادة المبيعات والإيرادات وتحسين رضا العملاء. يتم استخدامها على نطاق واسع في مجموعة متنوعة من الصناعات، بما في ذلك التجارة الإلكترونية والإعلام والرعاية الصحية، ومن المرجح أن تصبح أكثر أهمية في المستقبل مع استمرار تقدم البيانات والتكنولوجيا.

## أمثلة على موقع الأفلام التي تحتوي على أنظمة توصية

1. **Netflix**: واحدة من أكثر خدمات بث الأفلام شيوعاً حيث تقدم الأفلام والأفلام الوثائقية والبرامج التلفزيونية. تحتوي على مجموعة كبيرة من المحتوى الأصلي، بما في ذلك العروض التي نالت استحساناً كبيراً مثل *The Stranger Things* و *Crown*. وهي متوفرة في أكثر من 190 دولة ولها رسوم اشتراك شهرية تختلف حسب الخطة التي تختارها. وهي معروفة بواجهة سهلة الاستخدام وتوصيات متخصصة. لا تقدم البث التلفزيوني المباشر.

2. **Hulu**: هي خدمة بث مقرها الولايات المتحدة تقدم برامج تلفزيونية وأفلاماً وقنوات تلفزيونية مباشرة. تحتوي على مجموعة كبيرة من البرامج التلفزيونية الحالية

والسابقة، بما في ذلك بعض برامج الشبكة الأكثر شهرة This Is Us و Bachelor. كما أن لديها مجموعة مختارة من الأفلام، على الرغم من أنها ليست واسعة النطاق مثل Netflix. لها رسوم شهرية مع الإعلانات أو بدون الإعلانات. وهي معروفة بشكل خاص بتقديم وصول إلى حلقات العرض القادمة من البرامج التلفزيونية الشهيرة.

**Amazon Prime Video .3**: تعد Amazon Prime Video جزءاً من خدمة اشتراك Amazon Prime، والتي تشمل أيضاً الشحن المجاني والموسيقى والامتيازات الأخرى تقدم مجموعة مختارة من الأفلام والبرامج التلفزيونية، بما في ذلك إنتاج أمازون الأصلي مثل Mrs. Maisel و The Marvelous Mrs. Maisel. متوفراً في العديد من البلدان ولها رسوم اشتراك شهرية تختلف حسب البلد. وهي معروفة بشكل خاص بمجموعتها الكبيرة من الأفلام الكلاسيكية وواجهة سهلة الاستخدام.

**HBO Max .4**: هي خدمة بث تقدم مجموعة متنوعة من الأفلام والبرامج التلفزيونية، بما في ذلك أفلام HBO الأصلية مثل The Game of Thrones و The Sopranos. كما أن لديها مجموعة مختارة من الأفلام، على الرغم من أنها ليست واسعة النطاق مثل Netflix. متاحة فقط في الولايات المتحدة ولها رسوم اشتراك شهرية.

**Disney+ .5**: هي خدمة بث تركز على المحتوى المناسب للعائلة والمحتوى الحصري من Disney و Pixar و Marvel و Star Wars. تحتوي على مجموعة كبيرة من الأفلام والبرامج التلفزيونية، بما في ذلك العديد من الكلاسيكيات مثل The Snow White and the Seven Dwarfs وأحدث الأغاني مثل The Mandalorian. وهي متوفرة في العديد من البلدان ولها رسوم اشتراك شهرية.

**Apple TV+ .6**: هي خدمة بث تقدم مزيجاً من المحتوى الأصلي. وهي معروفة بشكل خاص بإنتاجاتها عالية الجودة وتركيزها على الجودة أكثر من الكمية.

**Peacock .7**: هي خدمة بث تقدم مجموعة متنوعة من الأفلام والبرامج التلفزيونية، بما في ذلك عروض NBCUniversal الأصلية. وهي معروفة بشكل خاص بالبث المباشر للرياضيات وطبقة الإعلانات المجانية المدعومة.

8. **Tubi** : هي خدمة بث مجانية مدعاومة بالإعلانات تقدم مجموعة مختارة من الأفلام والبرامج التلفزيونية. تشتهر بشكل خاص بمجموعتها الكبيرة من الأفلام الكلاسيكية وواجهة الاستخدام.

9. **Crave** : خدمة بث كندية تقدم مزيجاً من البرامج التلفزيونية والأفلام والمحظى الأصلي. وهي معروفة بشكل خاص بوصولها الحصري إلى محتوى HBO في كندا.

10. **Vudu** : هي خدمة بث تقدم مزيجاً من الأفلام والبرامج التلفزيونية للشراء أو الاستئجار. وهي معروفة بشكل خاص بجودة الفيديو والصوت عالية الجودة وقدرتها على بث المحتوى على مجموعة واسعة من الأجهزة.

## مشكلة البحث

الانتشار الكبير لمنصات عرض الأفلام والمسلسلات التي قد لا تلبي طلب المشاهد من خلال إظهار خيارات غير ذات صلة بما يفضل المستخدم او ما يبحث عنه او ما يمكن أن يجذب انتباذه.

## الهدف

إن تطبيق نظام توصية فعال على نظام ما – منصات عرض الأفلام في مشروعنا – يهدف إلى إضفاء الذكاء على توصيات النظام وبالتالي إرضاء المشاهد وزيادة عدد المشاهدات والتقييمات للمحتوى مما يؤدي إلى زيادة في الأرباح.

## **النتائج المتوقعة**

بناء نظام توصية فعال قادر على تلبية تطلعات المستخدم وإظهار الخيارات الأفضل له دوماً، ويستفيد منه جميع الأشخاص بمختلف الأعمار والأجناس التي تفضل مشاهدة الأفلام والمسلسلات في أوقات فراغهم.

## **ترتيب الأطروحة**

**الفصل الأول:** يقوم بشرح بعض الدراسات النظرية.

**الفصل الثاني:** يقوم بطرح بعض التطبيقات المتعلقة.

**الفصل الثالث:** القسم العملي

**الفصل الرابع:** النتائج

**الفصل الخامس:** الآفاق المستقبلية

## **1- الفصل الأول: الدراسة النظرية**

### **1-1 دراسة حول موقع الأفلام ومتطلباتها**

بناءً على التنوع الكبير في موقع الأفلام والأعداد الهائلة للمستخدمين في العصر الحالي، يتبع علينا إقامة معايير قياسية تميز بين هذه المواقع. سنقوم في هذه الدراسة بفحص هذا الموضوع من منظوريين: وجهة نظر المستخدم العادي ووجهة نظر المبرمج

#### **1-1-1 وجهة نظر المستخدم**

1. اختيار المحتوى: تنوع وجودة الأفلام يلعبان دوراً حاسماً في جذب المستخدمين لاستخدام الموقع.
2. واجهة المستخدم: صميم جذاب وبيهقي لواجهة المستخدم يعزز تجربة المستخدم، مع سهولة الاستخدام وعرض واضح للمعلومات.
3. خاصية البحث: البحث الفعال يساعد المستخدمين في العثور على الأفلام بسرعة، مما يجعلها خاصية ضرورية.
4. نظام التوصية يساعد المستخدمين في اكتشاف أفلام جديدة بناءً على تفضيلاتهم وتاريخ مشاهدتهم.
5. جودة البث: البث عالي الجودة مع أقل تخزين مؤقت يعزز تجربة مشاهدة المستخدم.
6. ميزات التواصل الاجتماعي: القدرة على التواصل مع مستخدمين آخرين ومشاركة الآراء تعزز تفاعل المستخدم.
7. تطبيق الموبايل: توفر تطبيقات الموبايل وقابليتها للاستخدام تعزز التفاعل مع المحتوى.
8. التكلفة: تقديم رسوم اشتراك معقولة يعد عاملاً مهماً.
9. التوفير: يجب أن يكون المحتوى متاحاً في مناطق مختلفة من العالم.
10. الإعلانات: على الرغم من العوائد التي تتحققها تلك الإعلانات، إلا أنها تؤثر على تجربة المستخدم. ويعتمد هذا التأثير على نوع الإعلانات وموضعها.

## 1-1-2 من وجهة نظر المبرمج

1. التقنيات المستخدمة: استخدام لغات البرمجة وأطر العمل والمكتبات الحديثة والموثوقة يمكن أن يسهم في تحسين أداء الموقع وتعزيز أمانه وقابليته للتوسيع.
2. تصميم واجهة المستخدم: تصميم واجهة المستخدم بشكل جذاب وسهل التنقل يعزز بشكل كبير تجربة المستخدم، مع توفير ميزات مفيدة مثل البحث والفلاتر والتوصيات.
3. إدارة قواعد البيانات: قدرة الموقع على تخزين واسترداد كميات كبيرة من البيانات بكفاءة تعتبر أمراً حيوياً، بما في ذلك تفاصيل الفيلم ومراجعات المستخدمين والمعلومات ذات الصلة.
4. اختيار المحتوى: يلعب نظام إدارة المحتوى الفعال دوراً رئيسياً في توفير تنوع وجودة الأفلام على الموقع، مع إمكانية التحديث السهل لمعلومات الفيلم والمقاطع الدعائية والتقييمات.
5. التكامل مع خدمات الجهات الخارجية: يتطلب التكامل الفعال مع خدمات الجهات الخارجية، مثل بوابات الدفع ومنصات البث وقنوات التواصل الاجتماعي، لتحسين تجربة المستخدم وتوسيع إمكانيات الموقع.
6. قابلية التوسيع: يجب أن يكون الموقع قادراً على التعامل مع حجم كبير من المستخدمين وحركة المرور بفعالية ودون مشاكل في الأداء.
7. الأمان: يعتبر الأمان أمراً حاسماً، حيث يجب أن يتبع الموقع أحدث بروتوكولات وتقنيات الأمان لحماية المعلومات الشخصية والمالية للمستخدمين.
8. الأداء: يتطلب الأداء الجيد تحسين سرعة الموقع وكفاءته، مما يتيح للمستخدمين تجربة تصفح سلسة وتحميل سريع.
9. تكامل واجهة برمجة التطبيقات API: يسمح تكامل واجهات برمجة التطبيقات بالتفاعل مع خدمات أخرى مثل موقع التواصل الاجتماعي وبوابات الدفع ومحركات التوصية.
10. إمكانية الوصول: يجب على الموقع الامتثال لمعايير إمكانية الوصول لدعم جميع فئات المستخدمين، بما في ذلك دعم برامج القراءة الشاشة وتقنيات المساعدة الأخرى.

11. استجابة الهاتف المحمول: يجب أن يكون الموقع قابلاً للاستجابة وقدراً على التكيف مع مختلف أحجام الشاشات وأنواع الأجهزة، مثل الهواتف الذكية والأجهزة اللوحية، لتحسين تجربة المستخدم.

سلط هذه الدراسة الضوء على أهمية نظام التوصية في تحسين تجربة المستخدم على موقع الأفلام. يُشدد على الجوانب الفنية الأساسية التي يجب أن تكون متوفرة للمستخدمين والمبرمجين لضمان نجاح مثل هذه المنصات. يثير الاهتمام الفعال بنظام التوصية تساؤلات حول طبيعته، ولماذا يعتبر عنصراً رئيسياً لنجاح أي موقع مخصص للأفلام، بالإضافة إلى استكشاف أنواع مختلفة من أنظمة التوصية وفهم الاختلافات بينها.(2)

## 1-2-تعريف نظام التوصية

هو نوع من خوارزميات الذكاء الصنعي التي تحلل بيانات المستخدم وسلوكه وتفضيلاته لاقتراح العناصر أو المحتوى أو الإجراءات التي قد تهم المستخدم. تستخدم هذه الأنظمة بشكل شائع في التجارة الإلكترونية ووسائل التواصل الاجتماعي وغيرها من المنصات عبر الانترنت لتقديم توصيات مخصصة للمستخدمين، بناء على سلوكهم السابق وسلوك المستخدمين الآخرين المماثلين.

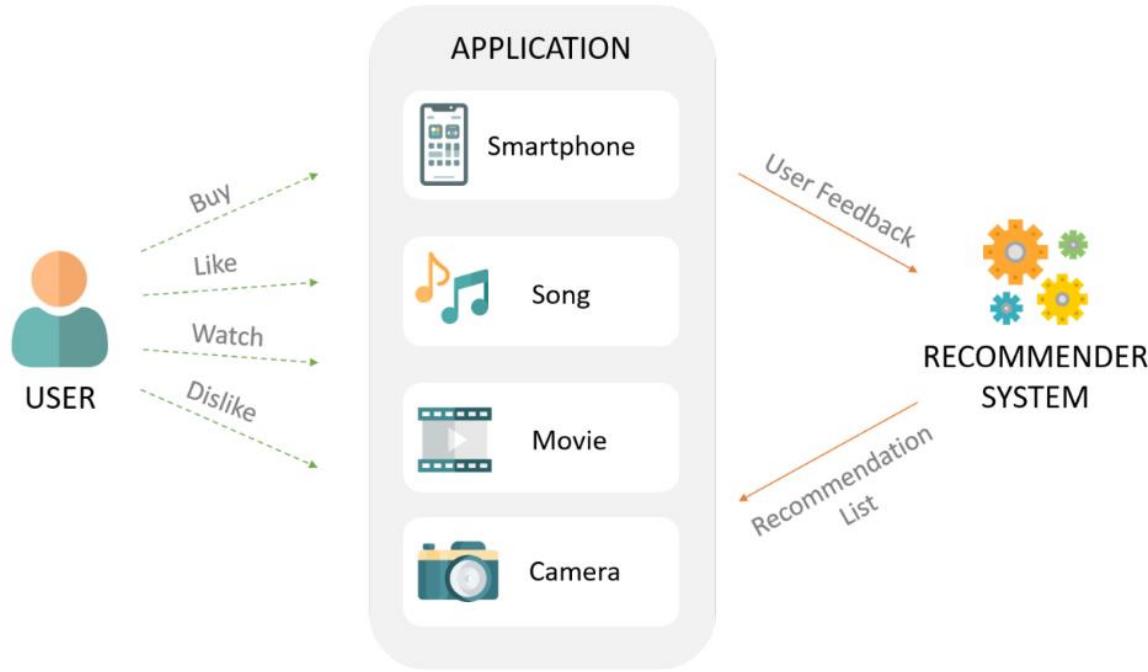
وبالتالي تم تصميم أنظمة التوصية لتقدير فائدة عنصر ما والتئي بـما إذا كان يستحق التوصية. العنصر الأساسي لنظام التوصية هو: (3)

$$f: U * I \rightarrow D.$$

هذه دالة لتحديد فائدة عنصر معين  $u \in U$  للمستخدم  $i \in I$ . حيث أن  $D$  هي قائمة التوصيات النهائية التي تحتوي على مجموعة من العناصر مرتبة وفقاً لفائدة جميع العناصر التي لم يستهلكها المستخدم [11]. يتم تقديم فائدة عنصر ما من حيث تصنيفات المستخدم. تجد أنظمة التوصية عنصراً للمستخدم من خلال زيادة دالة المنفعة إلى الحد الأقصى على النحو التالي:

$$\forall u \in U, \text{argmax}_{i \in I} f(u, i)$$

يختلف توقع فائدة العناصر المستخدم معين وفقاً لخوارزمية التوصية المحددة.



الشكل 2: خوارزميات التوصية

هناك نوعان رئيسيان من أنظمة التوصيات: التصفية التعاونية Collaborative Filtering والتصفية القائمة على المحتوى Content-based Filtering.

توصي خوارزميات التصفية التعاونية بالعناصر بناء على تشابه سلوكيات المستخدمين وتفضيلاتهم.

توصي خوارزميات التصفية القائمة على المحتوى بالعناصر بناء على تشابه ميزات العناصر وتفضيلات المستخدم.

هناك أيضاً أنظمة توصية مختلطة تجمع بين تقنيات التصفية التعاونية والقائمة على المحتوى لتقديم توصيات أكثر دقة وتنوعاً.

بالإضافة إلى هذه الأنواع الأساسية، هناك أيضاً أنظمة توصية أكثر تقدماً تستخدم تقنيات مثل عامل المصفوفة والتعلم العميق والتعليم المعزز لتحسين دقة التوصيات وأهميتها.

سنتطرق إلى أنواع أنظمة التوصية المختلفة في الفقرات اللاحقة.

## 1-3- أنظمة التوصية

### 1-3-1 أنواع أنظمة التوصية

توجد العديد من أنواع أنظمة التوصية، ومن بينها: (4)

I. التصفية التعاونية **Collaborative Filtering**: التصفية التعاونية هي تقنية تحل سلوك المستخدم وتفضيلاته للتوصية بالعناصر المشابهة لتلك التي أعجبتهم في الماضي بالإضافة إلى تفضيلات المستخدمين المشابهين للمستخدم الحالي. يعمل هذا النوع من أنظمة التوصية بشكل فعال عندما يكون هناك الكثير من البيانات المتاحة حول سلوك المستخدم، ولكن يمكن أن يواجه مشكلة "البداية الباردة" حيث لا يتتوفر سوى القليل من البيانات حول المستخدمين الجدد.

#### كيفية عمل Collaborative Filtering

1) جمع البيانات: يتم جمع بيانات التقييمات والتفضيلات من المستخدمين الذين قاموا

بتقييم العناصر. يمكن تمثيل هذه البيانات في مصفوفة  $R$  حيث يكون  $R_{ij}$  هو تقييم

المستخدم  $i$  للعنصر  $j$

2) بناء المصفوفة: يتم بناء مصفوفة التفضيلات والتقييمات  $R$ ، حيث تمثل الصفوف

المستخدمين والأعمدة العناصر

	item 1	item 2	item 3	...	item n
user 1					
user 2					
user 3					
user 4					
user 5					
user 6					
user 7					
user 8					
...					
user n					

$\underline{R}$

الشكل 3: المصفوفة  $R$

(3) حساب التشابه: يتم حساب درجة التشابه بين المستخدم الحالي  $u$  و مجموعة من المستخدمين الآخرين. يمكن تحديد درجة التشابه باستخدام عدة مقاييس وأحد أهمها هو مقياس جيب التمام Cosine Similarity، وهي تُعَرِّف عن التشابه الزاوي بين متوجه التفضيل:

$$\text{Cosine Similarity}(u, v) = \frac{\sum_i R_{ui} R_{vi}}{\sqrt{\sum_i R_{ui}} \cdot \sqrt{\sum_i R_{vi}}}$$

(4) توليد التوصية: يتم توليد التوصية عن طريق تقييم تفضيلات المستخدم وتوصية العناصر المشابهة التي تم تقييمها بواسطة المستخدمين الآخرين. بعد حساب درجات التشابه، يمكن توجيه التوصيات باستخدام وزن مرتبط بدرجة التشابه. يمكن استخدام معادلة التوصية كالتالي:

$$\hat{R}_{ui} = \frac{\sum_v \text{Similarity}(u, v) \cdot R_{vi}}{\sum_v |\text{Similarity}(u, v)|}$$

### **:Collaborative Filtering تطبيقات**

- **منصات البث المباشر:** يمكن تطبيق نظام Collaborative Filtering في منصات البث المباشر لتوصية المستخدمين بمحتوى ذو اهتمام استناداً إلى تقييمات المستخدمين الآخرين.
- **موقع التجارة الإلكترونية:** يستخدم لتوصية المستخدمين بالمنتجات المشابهة لتلك التي قاموا بشرائها أو تقييمها.
- **تطبيقات التواصل الاجتماعي:** يمكن استخدامه لتوصية المستخدمين بالأصدقاء أو المجموعات المشابهة باستناد إلى تفضيلاتهم.

## **مزايا Collaborative Filtering**

1. الاعتماد على البيانات الفعلية: تعتمد التصفية التعاونية على بيانات التقييمات والتفضيلات الفعلية التي يقدمها المستخدمون، مما يزيد من دقة التوصيات عندما تكون البيانات كافية.
2. تطبيق متعدد المجالات: يمكن تطبيقها بسهولة في مجالات متعددة مثل الموسيقى، الأفلام، الكتب، والمنتجات الأخرى، مما يجعلها مرنّة وقابلة للاستخدام في سياقات مختلفة.

## **عيوب Collaborative Filtering**

1. قد تكون غير دقيقة في حالة عدم كفاية البيانات: إذا كان هناك قلة في بيانات التقييم لعدد كبير من المستخدمين، قد تتراجع دقة التصفية التعاونية.
2. الحاجة إلى تفاعل كبير من المستخدمين: يتطلب نظام التصفية التعاونية التفاعل الكبير بين المستخدمين لتحقيق دقة عالية في التوصيات.
3. التأثر بالتغييرات في تفضيلات المستخدمين وتقييماتهم: قد يكون التأثر بغيرات في تفضيلات المستخدمين وتقييماتهم تحدياً، حيث يتغير سلوك المستخدمين بشكل مستمر.

## **معلومات هامة حول نظام التوصية :Collaborative Filtering**

- هناك نوعان من Collaborative Filtering، الأول يسمى User-based، ويعتمد على مقارنة تفضيلات المستخدم الحالي مع تفضيلات مجموعة من المستخدمين المشابهين له، والنوع الثاني يسمى Item-based، ويعتمد على مقارنة التفضيلات بين العناصر المختلفة وتوصية العناصر المشابهة لتلك التي قام المستخدم بتقييمها.
- يمكن تحسين أداء نظام التوصية Collaborative Filtering عن طريق استخدام تقنيات مثل Singular Value Decomposition (SVD) و Factorization Matrix وغيرها.

- يمكن أن يعاني Collaborative Filtering من مشكلة البداية الباردة "The Cold Start" ، وهي عدم وجود بيانات تقييمات كافية لبعض المستخدمين الجدد أو المنتجات الجديدة، مما يؤدي إلى صعوبة توصية عناصر مشابهة لتلك العناصر.
  - يمكن دمج نظام التوصية Collaborative Filtering مع نظام التصفية القائم على المحتوى Content-Based Filtering لتحسين الدقة والاستفادة من مزايا كل نظام.
  - يمكن تطبيق Collaborative Filtering على المجموعات الصغيرة من المستخدمين والعناصر، المعروفة باسم "Small Data" ، وذلك لأنه يستند إلى التفضيلات الفردية والتي يمكن جمعها بشكل متاح ومناسب.
- II. التصفية القائمة على المحتوى Content-based Filtering: تُعد التصفية القائمة على المحتوى إحدى تقنيات نظم التوصية، حيث يتم توجيه المستخدمين نحو عناصر معينة استناداً إلى خصائصها وميزاتها. يتأسس أداء هذه الأنظمة على فعاليتها في حالة توافر كميات كبيرة من البيانات حول العناصر الموصى بها. ومع ذلك، قد تعاني هذه الأنظمة في توجيه المستخدمين نحو عناصر جديدة ومتعددة يمكن أن تثير اهتمامهم. على سبيل المثال، إذا كان المستخدم قد أبدى تفضيلاً مسبقاً لأفلام الحركة، فإن النظام قد يقترح أفلام حركة أخرى. يعتمد هذا الأسلوب على استخدام خوارزميات مثل تشابه جيب التمام Cosine Similarity أو (تردد المصطلح - التردد العكسي للمستند) TF-IDF لمقارنة سمات العناصر الموصى بها بتلك التي تم الإعجاب بها في الماضي.

### كيفية عمل Content-based Filtering

1) جمع البيانات: يتطلب تنفيذ نظام التصفية القائمة على المحتوى جمع بيانات شاملة حول المنتجات المتاحة والمعلومات الوصفية للعناصر. يمكن تمثيل هذه البيانات في مصفوفة  $R$  حيث يكون  $R_{ij}$  هو السمة  $j$  العنصر أو المستخدم  $i$

2) تحليل المحتوى: يجري تحليل محتوى العناصر للبحث عن معلومات وصفية مشابهة لتلك التي يفضلها المستخدم. وذلك من خلال حساب درجة التشابه بين المحتوى الموجود وتفضيلات المستخدم. ويتم ذلك من خلال استخدام أحد الخوارزميات الشهيرة مثل جيب التمام Cosine Similarity التي شرحناها سابقاً

في نظام التصفية التعاونية. أو كمثال آخر TF-IDF (تردد المصطلح - التردد العكسي للمستند) التي تُستخدم لتقدير أهمية كلمة معينة في وثيقة أو بناءً على مدى تكرارها واهتمامها في النص بشكل عام. يمكن استخدامها لتمثيل المحتوى بشكل رقمي، وبالتالي، يمكن استخدامها في حساب تشابه المحتوى بين المنتجات وتفضيلات المستخدم:

1. معادلة **TF**: يقاس تكرار كلمة معينة في وثيقة ما وتعطى بالمعادلة التالية:

$$TF(t, d) = \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d}$$

2. معادلة **IDF**: يقاس بمدى استمرارية كلمة عبر مجموعة من الوثائق وتعطى بالمعادلة التالية:

$$IDF(t, D) = \log \left( \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d} \right)$$

حيث  $|D|$  هو إجمالي عدد الوثائق

3. معادلة **TF-IDF**: تجمع بين معادلتي TF و IDF لتحديد وزن الكلمة في وثيقة معينة في سياق المجموعة وتعطى بالمعادلة التالية:

$$TF - IDF(t, d, D) = TF(t, d) \times IDF(t, D)$$

(3) توليد التوصية: يتم توليد التوصية عن طريق مقارنة السمات المستخرجة ( $S$ ) مع تفضيلات المستخدم. وتحديد العناصر المشابهة لفضيلاته.

### تطبيقات Content-based Filtering

يستخدم نظام التصفية القائم على المحتوى في مجموعة من التطبيقات مثل:

- تطبيقات الأخبار والمقالات لتوصية المستخدمين بالمقالات ذات الصلة.

- منصات البث المباشر لتوصية المستخدمين بالأفلام والبرامج التلفزيونية المشابهة لتلك التي يفضلونها.
- موقع التجارة الإلكترونية لتوصية المستخدمين بالمنتجات المشابهة لتلك التي يفضلونها.

### **:Content-based Filtering مزايا**

1. الاعتماد على معلومات المحتوى الفعلية: يعتمد النظام على محتوى العناصر بشكل مباشر، مما يعني استخدام معلومات حقيقة وواقعية حول العناصر لتحديد التوصيات.
2. تطبيق متعدد المجالات: يمكن تكامل نظام التوصية في مجالات متعددة مثل الأفلام والكتب والمنتجات الإلكترونية، مما يسهم في توسيع نطاق تطبيقه في مختلف الصناعات.

### **:Content-based Filtering عيوب**

1. قد تكون غير دقيقة في حالة التشابه: في حالة وجود بيانات تفضيلات مشابهة لدى مجموعة كبيرة من المستخدمين، قد يكون النظام غير دقيق في تحديد التوصيات الفريدة لكل مستخدم.
2. الاعتماد على المعلومات الوصفية: يعتمد النظام بشكل رئيسي على المعلومات الوصفية للعناصر، مما يعني أنه قد لا يأخذ بعين الاعتبار العوامل الاجتماعية والنفسية والشخصية للمستخدمين.
3. تكرار التوصيات: قد ينتج النظام عن تكرار التوصيات لنفس العناصر أو الأنواع المشابهة، مما يمكن أن يؤدي إلى قلة التنوع في التوصيات.
4. تأثر بالتغييرات: قد يتأثر أداء النظام بشكل كبير عندما يتغير نوع المحتوى أو المجال الذي يتم توصية المستخدمين بشأنه.

## معلومات هامة حول نظام التوصية :Content-based Filtering

- يتطلب نظام التوصية Content-based Filtering الحصول على معلومات وصفية دقيقة وشاملة للعناصر، وهذا يمكن أن يكون تحدياً في بعض المجالات مثل الموسيقى والفنون.
- يمكن تحسين أداء نظام التوصية Content-based Filtering عن طريق استخدام تقنيات مثل Machine Learning و Natural Language Processing (NLP) وغيرها.
- يمكن دمج نظام التوصية Content-based Filtering مع نظام التصفية القائم على المستخدم User-based Filtering لتحسين الدقة والاستفادة من مزايا كل نظام.
- يمكن أن يتأثر أداء نظام التوصية Content-based Filtering بشكل كبير عندما يتغير نوع المحتوى أو المجال الذي يتم توصية المستخدمين بشأنه.
- يمكن تطبيق نظام التوصية Content-based Filtering على المجموعات الصغيرة من المستخدمين والعناصر، المعروفة باسم "Small Data" ، وذلك لأنها تعتمد على المعلومات الوصفية للعناصر والتي يمكن جمعها بشكل متاح ومناسب.
- يمكن أن يستخدم نظام التوصية Content-based Filtering بشكل فعال في مجالات مثل التسويق الإلكتروني والإعلانات المستهدفة، حيث يمكن استخدام المعلومات الوصفية للعناصر لتحديد المستهلكين المستهدفين بشكل أفضل.

III. **أنظمة التوصية الهجينية Hybrid Recommendation Systems:** أنظمة التوصية الهجينية تمثل تقدماً مهماً في مجال أنظمة التوصية، حيث تدمج بين نوعين أو أكثر من أساليب التوصية لتحقيق توصيات أكثر دقة وكفاءة. يتيح هذا الدمج للنظام استفادة من مميزات كل نوع من أنواع التوصية والتغلب على العيوب المحتملة. على سبيل المثال، يمكن لنظام التوصية الهجين الجمع بين التصفية التعاونية والتصفية القائمة على المحتوى لتحسين جودة التوصيات.

## **كيفية عمل Hybrid Recommendation Systems**

- 1) **جمع البيانات:** يتم جمع بيانات شاملة حول المستخدمين والعناصر المتوفرة بالإضافة إلى المعلومات الوصفية للعناصر.
- 2) **تحليل المحتوى:** يتم تحليل المحتوى للعناصر، حيث يتم البحث عن المعلومات الوصفية المشابهة تلائم تفضيلات المستخدم.
- 3) **تحليل التفضيلات:** يتم تحليل تفضيلات المستخدمين والبحث عن مستخدمين ذوي تفضيلات مشابهة.
- 4) **توليد التوصية:** يتم توليد التوصية عبر دمج توصيات نظامي التصفية التعاونية والتصفية القائمة على المحتوى.

## **تطبيقات Hybrid Recommendation Systems**

تستخدم أنظمة التوصية الهجينة في مجموعة من التطبيقات مثل:

- يستخدم في منصات البث المباشر لتوصية المستخدمين بالبرامج المشابهة وفي الوقت نفسه يقدم توصيات استناداً إلى تفضيلات مستخدمين آخرين.
- يعتبر فعالاً في موقع التجارة الإلكترونية لتوصية المستخدمين بالمنتجات المشابهة وتوفير توصيات مبنية على خصائص مماثلة.

## **مزايا Hybrid Recommendation Systems**

1. **تحسين دقة التوصية:** يمكن تحسين دقة التوصية عن طريق استخدام مزيج من تقنيات التوصية المختلفة.
2. **تفادي عيوب النظم الفردية:** يسمح بتفادي عيوب كل نظام فردي من خلال استغلال مزايا نظامي التصفية التعاونية والتصفية القائمة على المحتوى.

## **عيوب Hybrid Recommendation Systems**

1. **تعقيد في التنفيذ:** قد يكون تنفيذ نظام التوصية الهجين أكثر تعقيداً من تطبيق كل نظام على حدة.

2. **صعوبة اختيار المزيج المناسب:** قد يكون صعباً اختيار المزيج المثلى لتحقيق توازن بين التصفية التعاونية والتصفية القائمة على المحتوى.

### معلومات هامة حول نظام التوصية :**Hybrid Recommendation Systems**

- يمكن تطبيق نظام التوصية Hybrid Filtering بعدة طرق مختلفة، مثل الجمع بين توصيات نظامي التوصية Content-based و Collaborative Filtering أو الاعتماد على نظامي التوصية بشكل متكامل ومتتابع Filtering.
- يمكن تحسين أداء نظام التوصية Hybrid Filtering عن طريق استخدام تقنيات مثل Deep Learning وغيرها من التقنيات المتقدمة.
- يمكن استخدام نظام التوصية Hybrid Filtering في تحسين تجربة المستخدم في العديد من المجالات، مثل التسويق الإلكتروني والتجارة الإلكترونية والبث المباشر والمزيد.
- يمكن تطبيق نظام التوصية Hybrid Filtering على المجموعات الصغيرة من المستخدمين والعناصر، المعروفة باسم "Small Data"، وذلك لأنه يعتمد على المعلومات الوصفية للعناصر وتقديرات المستخدمين التي يمكن جمعها بشكل متاح ومناسب.
- يمكن أن يساعد نظام التوصية Hybrid Filtering على تحسين التنوع في التوصيات وتقديم خيارات مختلفة للمستخدمين، وذلك بسبب تطبيق مزيج من نظامي التوصية Content-based Filtering و Collaborative Filtering.
- يمكن استخدام نظام التوصية Hybrid Filtering لتحديد العوامل المؤثرة في تفضيلات المستخدمين وتحديد العوامل التي يمكن تحسينها لتحسين تجربة المستخدم.

## IV. أنظمة التوصية القائمة على المعرفة Knowledge-based recommendation

systems: تعتمد أنظمة التوصية القائمة على المعرفة على معرفة مسبقة حول المنتجات وفضائل المستخدمين لتقديم توصيات دقيقة. يمكن أن تكون هذه الأنظمة فعالة بشكل خاص عندما يكون لدى المستخدم تفضيلات أو قيود محددة بوضوح. ومع ذلك، قد تواجه صعوبة في توصية المستخدمين بالعناصر الجديدة والمتنوعة التي قد تناسب ذوقهم.

على سبيل المثال، يمكن لموقع السفر مثل Wego أن يستخدم نظاماً قائماً على المعرفة لتوجيه المستخدمين نحو الفنادق بناءً على ميزانيتهم وموقعهم المفضل ووسائل الراحة التي يفضلونها. هنا يتم تحليل معرفة سابقة عن الفنادق وتحديد توافقها مع احتياجات المستخدم

### كيفية عمل Knowledge-based recommendation systems

1) تحليل المعرفة المسبقة: يتم تحليل المعرفة المسبقة حول المنتجات وفضائلها واحتياجات المستخدمين.

2) تحليل الاحتياجات: يتم تحليل احتياجات المستخدم وتحديد المنتجات التي تلبي تلك الاحتياجات.

3) تحليل المنتجات: يتم تحليل المنتجات وتحديد توافقها مع تفضيلات المستخدم.

4) توليد التوصية: يتم توليد التوصية عن طريق توجيه المستخدم نحو المنتجات التي تتناسب مع تفضيلاته.

### تطبيقات Knowledge-based recommendation systems

يستخدم نظام التوصية القائم على المعرفة في مجموعة من التطبيقات مثل:

- يستخدم في التجارة الإلكترونية لتوجيه المستخدمين إلى المنتجات الملائمة لهم.
- يُستخدم في تطبيقات السفر والسياحة لتوصية المستخدمين بالوجهات والأنشطة التي تناسب اهتماماتهم.

### مزایا Knowledge-based recommendation systems

1. استخدام المعرفة المسبقة: يحسن استخدام المعرفة المسبقة دقة التوصية.

2. **فعالية في المنتجات الجديدة:** يمكن أن يكون فعالاً في توصية المستخدمين بالمنتجات الجديدة أو المتخصصة.

### **عيوب Knowledge-based recommendation systems**

1. **قلة التنوع:** قد يكون أقل فعالية في توصية المستخدمين بخيارات مختلفة.
2. **تكلفة التحديث:** يتطلب المعرفة المسبقة المناسبة حول المنتجات والخدمات واحتياجات المستخدمين، وهذا يمكن أن يكون مكلفاً وصعباً في بعض الأحيان.
3. **التحدي في تحليل تفضيلات المستخدم:** يمكن أن يكون تحديد احتياجات المستخدمين وحاجاتهم متغيراً بشكل مستمر، وهذا يعني أن نظام التصفية القائم على المحتوى يحتاج إلى تحديث مستمر لتلبية احتياجات المستخدمين.

### **معلومات هامة حول نظام التوصية Knowledge-based recommendation systems**

- يستخدم نظام التوصية Knowledge-based Recommender Systems الخوارزميات المنطقية والاستنتاجية لتحليل المعرفة المسبقة حول المنتجات والخدمات واحتياجات المستخدمين وتوليد التوصيات.
- يمكن استخدام نظام التوصية Knowledge-based Recommender Systems لتحليل وتصنيف المنتجات والخدمات وفقاً لميزاتها واحتياجات المستخدمين، مما يسهل على المستخدمين العثور على المنتجات التي يحتاجونها بسهولة.
- يمكن استخدام نظام التوصية Knowledge-based Recommender Systems في تحليل الاحتياجات الفردية للمستخدمين وتوصية المنتجات والخدمات التي تناسب احتياجاتهم، مما يساعد على تحسين تجربة المستخدم ورفع مستوى الرضا لديهم.
- يمكن تحسين أداء نظام التوصية Knowledge-based Recommender Systems عبر تحسين المعرفة المسبقة حول المنتجات والخدمات واحتياجات المستخدمين، وذلك عن طريق تحديث المعلومات المتاحة بشكل مستمر.

- يمكن استخدام نظام التوصية Knowledge-based Recommender Systems في تحليل السلوكيات السابقة للمستخدمين وتحليل تفضيلاتهم واحتياجاتهم وتوصية المنتجات والخدمات التي تتناسب معها.
- يمكن أن يوفر نظام التوصية Knowledge-based Recommender Systems توصيات شخصية ومخصصة لكل مستخدم، مما يعزز الارتباط بين المستخدم والموقع أو التطبيق ويسهل عليه البحث عن المنتجات التي يحتاجها.
- يعتبر نظام التوصية Knowledge-based Recommender Systems أقل فعالية من حيث دقة التوصية بالمقارنة مع النظم الأخرى مثل Collaborative Filtering، وذلك بسبب قلة المعرفة المسبقة المتاحة والتحديات المتعلقة بتحليل احتياجات المستخدمين.
- يتطلب نظام التوصية Knowledge-based Recommender Systems استخدام المعرفة المسبقة المناسبة لتوليد التوصيات، وهذا يمكن أن يتطلب تكلفة مرتفعة لجمع وتحليل المعلومات وتطوير نظام التوصية.

V. **أنظمة التوصية المدركة للسياق :Context Aware recommendation systems**

تعتمد أنظمة التوصية المدركة للسياق على مفهوم السياق الحالي، مثل الوقت، والمكان، والظروف الاجتماعية، لتحسين جودة التوصية للمستخدمين. فمثلاً، يمكن لتطبيق موسيقى مثل Anghami أن يوفر تجربة مخصصة بتوصيات تعتمد على الوقت من اليوم أو الطقس. تعتبر هذه الأنظمة فعالة عندما تتغير تفضيلات المستخدمين بناءً على السياق، ولكن يمكنها مواجهة تحدي "البداية الباردة" عند التعامل مع عناصر جديدة.

#### **كيفية عمل Context Aware recommendation systems**

- 1) **جمع السياق:** يتم جمع السياق الذي يتم فيه استخدام المنتجات أو الخدمات، مثل الوقت والمكان والعوامل الشخصية والظروف الاجتماعية.

2) **تحليل السياق:** يتم تحليل السياق وتحديد العوامل التي تؤثر على تفضيلات المستخدمين.

3) **تحليل المنتجات:** يتم تحليل المنتجات وتحديد توافقها مع السياق الذي يتم فيه استخدامها.

4) **توليد التوصية:** يتم توليد التوصية عن طريق توجيه المستخدم نحو المنتجات التي تناسب السياق الحالي.

### **:Context Aware recommendation systems تطبيقات**

يستخدم نظام التوصية المدرك للسياق في مجموعة من التطبيقات مثل:

- يستخدم في تطبيقات السفر والسياحة لتوصية المستخدمين بالأماكن والأنشطة المناسبة للسياق.
- يُستخدم في تطبيقات اللياقة البدنية لتوجيه المستخدمين نحو التمارين الملائمة للظروف الحالية.

### **:Context Aware recommendation systems مزايا**

1. **استخدام السياق:** يعتمد على السياق لتحسين دقة التوصية.
2. **فعالية في المواقف المختلفة:** يمكن أن يكون فعالاً في توصية المستخدمين بالمنتجات الملائمة للسياق الحالي.
3. **تحسين تجربة المستخدم:** يساعد في تحسين تجربة المستخدم ورفع مستوى الرضا لديهم.

### **:Context Aware recommendation systems عيوب**

1. **تحدي "البداية الباردة":** قد يكون أقل فعالية فيما يتعلق الأمر بالمنتجات الجديدة أو الغير مألوفة للسياق الذي يتم فيه استخدامها.

2. تكلفة جمع البيانات: يتطلب جمع وتحليل المعلومات الخاصة بالسياق الذي يتم فيه استخدام المنتجات أو الخدمات، وهذا يمكن أن يكون مكلفاً وصعباً في بعض الأحيان.

3. تحديث مستمر: يمكن أن يكون تحديد السياق الذي يتم فيه استخدام المنتجات أو الخدمات متغيراً بشكل مستمر، وهذا يعني أن نظام التوصية المدرك للسياق يحتاج إلى تحديث مستمر لتلبية احتياجات المستخدمين.

## معلومات هامة حول نظام التوصية Context Aware recommendation :systems

- يعتبر نظام التوصية Context-aware Recommender Systems أكثر فعالية في تحسين دقة التوصية بالمقارنة مع النظم الأخرى، حيث يأخذ في الاعتبار السياق الذي يتم فيه استخدام المنتجات أو الخدمات، وهذا يساعد على تقليل الأخطاء في التوصية.
- يمكن استخدام نظام التوصية Context-aware Recommender Systems في تحسين تجربة المستخدم وزيادة مستوى الرضا لديهم، إذ يمكن أن يتيح لهم التوصية بالمنتجات التي تناسب السياق الذي يتم فيه استخدامها.
- يمكن استخدام نظام التوصية Context-aware Recommender Systems في عدة تطبيقات مختلفة، مثل تطبيقات السفر والسياحة وتطبيقات الصحة واللياقة البدنية والتسوق عبر الإنترنت وغيرها، وذلك لتوصية المستخدمين بالمنتجات التي تناسب السياق الذي يتم فيه استخدامها.
- يمكن تحسين أداء نظام التوصية Context-aware Recommender Systems عبر تحديث المعلومات المتاحة بشكل مستمر، وذلك بجمع المزيد من البيانات حول السياق الذي يتم فيه استخدام المنتجات والخدمات وتحليلها لتحسين جودة التوصية.
- يمكن أن يكون نظام التوصية Context-aware Recommender Systems أقل فعالية في توصية المستخدمين بالمنتجات الجديدة أو الغير مألوفة للسياق الذي يتم فيه

استخدامها، إذ أنه يحتاج إلى مزيد من البيانات حول هذه المنتجات لتحسين جودة التوصية.

- يتطلب نظام التوصية Context-aware Recommender Systems جمع وتحليل المعلومات الخاصة بالسياق الذي يتم فيه استخدام المنتجات أو الخدمات، وهذا يمكن أن يكون مكلفاً وصعباً في بعض الأحيان، خاصة إذا كانت البيانات المتوفرة غير كافية.
- يمكن أن يكون تحديد السياق الذي يتم فيه استخدام المنتجات أو الخدمات متغيراً بشكل مستمر، وهذا يعني أن نظام التوصية Context-aware Recommender يتحتاج إلى تحديث مستمر لتلبية احتياجات المستخدمين Systems.

### 3-1-3-2 مقارنة بين أنظمة التوصية: نحو تحسين تجربة المستخدم في البيئة الرقمية

في هذه المقارنة، سنقوم بتسلیط الضوء على عدة أنظمة للتوصية للمستخدمين، وذلك بهدف فهم أفضل كيفية تحسين تجربة المستخدم في البيئة الرقمية. يلاحظ أن الأنظمة المطروحة تشمل Hybrid Filtering و Content-based Filtering و Collaborative Filtering و Context-aware Recommender و Knowledge-based Recommender Systems و Systems.

#### :Collaborative Filtering .1

- يعتمد على تقديم توصيات بناءً على تقييمات المستخدمين لمنتجات مشابهة.
- يتطلب كمية كبيرة من البيانات للتدريب والتنبؤ.
- يعاني من تحديد العوامل المؤثرة في توصية المستخدمين بدقة.

#### :Content-based Filtering .2

- يقوم بتوصية المستخدمين بمنتجات مشابهة لتلك التي قاموا بشرائها.

• يتطلب معرفة بمحتوى العناصر الموصى بها.

• قد يعاني من نقص المعلومات حول العناصر المختلفة.

### **:Hybrid Filtering .3**

• يتيح تحسين نظام الجودة بمزج ميزات نظامي توصية أو أكثر.

• يتطلب الكثير من البيانات للتدريب والتنبؤ.

• في مستويات معينة قد يكون معقداً.

### **:Knowledge-based Recommender Systems .4**

• يعتمد على المعرفة المتاحة حول المستخدمين والعناصر لتوصية المستخدمين.

• يستفيد من معرفة الملف الشخصي للمستخدمين والمعلومات حول العناصر.

• يمكن تطبيقه في الواقع الإلكتروني التي تحتاج إلى توصية المستخدمين بسبب عوامل معينة مثل العمر والجنس والاهتمامات.

### **:Context-aware Recommender Systems .5**

• يعتمد على معرفة السياق لتوصية العناصر المناسبة للمستخدمين.

• يستند إلى بيانات الاستخدام والتفاعل والملف الشخصي للمستخدمين والعناصر.

• يُعد خياراً فعالاً في توصية المستخدمين بالعناصر وفقاً للسياق الحالي.

و منه:

يعكس Collaborative Filtering أهمية تقييمات المستخدمين، ولكن يعاني من قضية كمية البيانات

Content-based Filtering يركز على تحليل المحتوى ويعتمد على معرفة دقيقة بالعناصر.

Hybrid Recommender Systems يجمع بين فوائد أنظمة التوصية المستخدمة ويضفي أهمية كبيرة في حال تجنب زيادة التعقيد في بناءه.

Knowledge-based Recommender Systems يعتمد على معرفة المستخدم والعناصر ويستند إلى ملف المستخدم.

Context-aware Recommender Systems يستفيد من معرفة السياق لتحسين دقة التوصية.

من الجدول (1) نستنتج ما يلي:

يمكن أن تختلف خصائص وفعالية أنظمة التوصية هذه من حيث مجالات تنفيذها، مجموعات البيانات وحالات الاستخدام. تعتمد النظم الهجينة على دمج التصفية التعاونية والتصفية القائمة إلى المحتوى للاستفادة من نقاط قوتهم وتقليل نقاط الضعف. التصفية المدركة للسياق تأخذ في اعتبارها معلومات السياق لتعزيز التوصيات. التصفية القائمة على المعرفة تعتمد على المعرفة الصريحة حول المستخدمين أو العناصر، غالباً ما يقدمها خبراء في المجال.

Context-Aware	Knowledge-Based	Hybrid	Content-Based Filtering	Collaborative Filtering	الخاصية
بيانات السياق الخاصة بالمستخدم، والعوامل البيئية	معرفة صريحة حول المستخدمين أو العناصر	مزيج من التصفية التعاونية والقائمة على المحتوى	سمات العناصر، ملف المستخدم	تفاعلات المستخدم مع العناصر	بيانات المدخلة
بيانات السياق الخاصة بالمستخدم، الموقع والوقت	معرفة صريحة حول المستخدمين أو العناصر	تفاعلات المستخدمين مع العناصر، وسمات المحتوى	سمات العناصر، تفضيلات المستخدم	تقييمات المستخدم للعناصر	متطلبات البيانات
أقل تواجداً	نعم، مالم تتوفر معرفة متخصصة	أقل تواجداً	نعم، ولكن أقل تواجداً من التصفية التعاونية	نعم	مشكلة البداية الباردة
تعتمد على تعقيد بيانات السياق	قابلة للتوسيع بشكل جيد مع بيانات المعرفة الصريحة	تعتمد على نهج النظام الهجين	جيءة مع مجموعة البيانات الكبيرة	شكل مشكلة مع مجموعات البيانات الكبيرة	قابلية للتوسيع
متوسطة إلى عالية	منخفضة مالم يتم تحديث بيانات المعرفة	تعتمد على نهج النظام الهجين	متوسطة	جيءة	الصدفة
متوسطة	عالية، بالاعتماد على القواعد الصريحة	تعتمد على نهج النظام الهجين	عالية	منخفضة	قابلية التفسير
متوسط إلى عالي	يعتمد على تمثيل المعرفة	يعتمد على نهج النظام الهجين	جيء	جيء	التعامل مع التنوع
متوسط إلى عالي	منخفض إلى متوسط	يعتمد على نهج النظام الهجين	منخفض إلى متوسط	متوسط إلى عالي	التعقيد
خدمات مبنية على الموقع، أخبار متخصصة.	الأنظمة الخبرية، التخخيص الطبي	نوصيات حيث يكون دمج نظم توصية مختلفة مقيداً	نوصيات الأفلام، توصيات المقالات	نوصيات الأفلام، توصيات المنتجات	حالات الاستخدام
يتكيف مع تغير السياق	يتطلب التحديثات لبيانات المعرفة	يمكن التكيف مع التغييرات في النظام الهجين	مرن للتغييرات المحتوى	تغير سلوك المستخدم مع مرور الوقت	المرونة
Google Now, Foursquare	الأنظمة الخبرية في الطب	Netflix (CF+CBF)	Pandora, Last.fm	Netflix, Amazon, Spotify	أمثلة

الجدول 1: مقارنة بين أنظمة التوصية المختلفة

### 3-1-3-3- استخدام أكثر من نظام توصية

يظهر التقدم المتزايد في تكنولوجيا نظم التوصية أهمية استخدام أكثر من نظام توصية في نفس الموقع الإلكتروني لتعزيز تجربة المستخدم وتلبية احتياجاته بفعالية. الاستخدام المترافق أو المترافق لمثل هذه الأنظمة يمكن أن يفتح آفاقاً جديدة لتحسين جودة التوصية.

إمكانية استخدام أنظمة متعددة:

Collaborative Filtering: يمكن استخدامه لتوصية المستخدمين بمنتجات مشابهة بناءً على تقييماتهم السابقة

Content-based Filtering: يستخدم لتوصية المستخدمين بمنتجات ذات محتوى مشابه أو مرتبطة بالمنتجات التي اشتروها.

Hybrid Filtering: يقدم ميزة جمع مزايا نظامي Collaborative Filtering و Content-based Filtering لتحسين دقة التوصية.

Knowledge-based Recommender Systems: يستخدم المعرفة المتوفرة المتوفرة لتجهيز المستخدمين نحو العناصر المناسبة استناداً إلى ملفاتهم الشخصية.

Context-aware Recommender Systems: يعزز التوصيات استناداً إلى السياق الحالي لاستخدام المستخدمين.

### فوائد الاستخدام المترافق

تحسين دقة التوصية: باستخدام مزيج من الأنظمة، يمكن تحسين دقة التوصية عن طريق اعتماد معلومات متنوعة.

تلبية احتياجات متنوعة للمستخدمين: يمكن لكل نظام توصية تلبية نمط محدد من احتياجات المستخدم، مما يزيد من شمولية التوصيات.

**توفر تجربة مستخدم شخصية:** يمكن تخصيص التوصيات بناءً على معلومات ملف المستخدم والبيئة.

### **العوامل الرئيسية لمراعاتها عند استخدام أكثر من نظام توصية**

1- **التجانس:** ضرورة التأكيد من أن الأنظمة المختلفة تتعامل مع نفس نوع البيانات، سواء كانت تقييمات المستخدمين أو بيانات المنتجات، لتحقيق التجانس وتسهيل الجمع والتحليل.

2- **المصداقية:** يجب التحقق من جودة ومصداقية البيانات المستخدمة في الأنظمة المختلفة، حيث يسهم ذلك في تفادي الأخطاء وتحسين دقة التوصية.

3- **الموثوقية:** يتعمّن التأكيد من أن الأنظمة تتمتع بمستوى عالي من الموثوقية، مما يسهم في حماية الأمان والخصوصية وتحسين جودة التوصية.

4- **الدقة:** من المهم مراقبة دقة الأنظمة بشكل دوري وتحديثها لتجنب الأخطاء وتحسين جودة التوصية بمرور الوقت.

5- **الكفاءة:** يجب تقييم كفاءة الأنظمة المختلفة وتأثيرها على أداء الموقع الإلكتروني بشكل عام لتحسين أداء الموقع وفعالية التوصية.

6- **التناسق:** يُفضل التأكيد من أن هناك توافقاً وتناسقاً بين الأنظمة المستخدمة والمحتوى المقدم، مما يعزز تجربة المستخدم بالتناغم والانسجام

يُشدد على أهمية مراعاة هذه العوامل أثناء استخدام أكثر من نظام توصية في نفس الموقع الإلكتروني. هذا يساهم في تحسين تجربة المستخدم وتقديم توصيات دقيقة وموثوقة، مع الحفاظ على معايير الجودة والأمان

### **4-3-1- سلبيات أنظمة التوصية**

بينما توفر أنظمة التوصية العديد من المزايا، إلا أن هناك أيضاً بعض العيوب المحتملة التي يجب مراعاتها، بما في ذلك: (5)

1. التنوع المحدود: قد تعاني أنظمة التوصية في بعض الأحيان من مشكلة "فقاعة التصنيفية"، حيث يتم تقديم توصيات للمستخدمين فقط تعزز تفضيلاتهم الحالية، مما يؤدي إلى نقص التنوع في المحتوى أو المنتجات التي يرونها.
  2. الاعتماد المفرط على السلوك السابق: قد تعتمد أنظمة التوصية بشكل كبير على سلوك المستخدم السابق وتفضيلاته، مما قد يؤدي إلى عرض محدود للمحتوى أو المنتجات الجديدة التي قد تكون ذات أهمية.
  3. مخاوف الخصوصية: تتطلب أنظمة التوصية الوصول إلى بيانات المستخدم من أجل تقديم توصيات مخصصة، والتي يمكن أن تثير مخاوف الخصوصية لبعض المستخدمين.
  4. التحيز الخوارزمي: قد تدخل أنظمة التوصية عن غير قصد تحيزات في التوصيات المقدمة، مما قد يؤدي إلى نتائج تمييزية أو غير عادلة.
  5. قابلية محدودة للتفسير: قد يكون من الصعب تفسير أو شرح بعض أنظمة التوصية، خاصة تلك التي تستخدم التعلم العميق أو غيرها من التقنيات المتقدمة، مما يجعل من الصعب على المستخدمين فهم سبب تقديم توصيات معينة.
- بشكل عام، في حين أن أنظمة التوصية يمكن أن تقدم العديد من الفوائد، إلا أن هناك أيضاً بعض الجوانب السلبية المحتملة التي يجب مراعاتها، بما في ذلك القضايا المتعلقة بالتنوع والخصوصية والتحيز والقابلية للتفسير.

## 2- الفصل الثاني: الدراسات المرجعية والأعمال المتعلقة

### 1-2- مراجعة الدراسات السابقة

تعتبر أنظمة التوصية مجالاً مهماً في الذكاء الاصطناعي وتعتبر واحدة من أهم التقنيات المستخدمة في التجارة الإلكترونية وصناعة المحتوى والترفيه وغيرها من المجالات. ولقد قام العديد من الباحثين والمؤلفين بإيجاد العديد من الأبحاث العلمية والمقالات المتعلقة بأنظمة التوصية وتحليل نتائجها والمساهمات التي قدمتها في المجال. ومن بين هذه المراجع:

"Deep Learning based Recommender System: A Survey and New -1

Hanwang Zhang و Lizi Liao و Xiangnan He، التي قدمها "Perspectives"

و Tat-Seng Chua و Liangyue Li، التي تتناول النظرة الشاملة

للتطورات الأخيرة في نظام التوصية القائم على التعلم العميق. (6)

Xiaoyuan، التي قدمها "A Survey of Collaborative Filtering Techniques" -2

و Taghi M. Khoshgoftaar و Su، التي تعرض للنظرة الشاملة على

تقنيات التصفية التعاونية (Collaborative Filtering). (7)

Dietmar Jannach، التي قدمها "Recommender Systems: An Introduction" -3

و Gerhard Friedrich و Alexander Felfernig و Markus Zanker، التي تتناول النظرة الشاملة لأنظمة التوصية والتحديات التي تواجهها والتقنيات

المستخدمة فيها. (8)

"Matrix Factorization Techniques for Recommender Systems" -4، التي

قدمها Chris Volinsky و Rober Bell و Koren Yehuda في عام 2009، والتي

تتناول النظرة الشاملة لتقنيات فصل العوامل Matrix Factorization في نظام التوصية.

(9)

"Hybrid Recommender Systems: Survey and Experiments" -5، التي قدمها

Robin Burke في عام 2002، والتي تتناول النظرة الشاملة على أنظمة التوصية الهجينية

والتحديات التي تواجهها. (4)

Christopher، وهو الكتاب الذي قدمه "Introduction to information retrieval" -6 Hinrich Schütze و Prabhakar Raghavan و D. Manning يقدم نظرة شاملة حول تقنيات استرجاع المعلومات، والذي يساهم في فهم وتنفيذ خوارزمية ال TF-IDF و Cosine Similarity في نظم التصفية القائمة على المحتوى لقياس التشابه بين العناصر. (10)

A Survey of Recommendation Systems: Recommendation Models, " -7 Hyeyoung Ko، قام بها الباحثون Techniques, and Application Fields Anna Choi و Yoonseo Park و Suyeon Lee بدراسة حول الاتجاهات البحثية التي تربط بين الجوانب التقنية المتقدمة لأنظمة التوصية المستخدمة في مختلف مجالات الخدمة والجوانب التجارية لهذه الخدمات. حيث قاموا بجمع ومراجعة أكثر من 135 مقالة من أعلى التصنيفات ومن أهم المؤتمرات المنشورة في Google Scholar بين عامي 2010 و 2021.

تتضمن المراجع السابقة العديد من المساهمات الهامة في مجال أنظمة التوصية، وتقدم نظرة شاملة حول الأنواع المختلفة لهذه الأنظمة وكيفية تحسين أدائها وجودتها. وقد أسهمت هذه الأبحاث والمقالات في تطوير وتحسين أنظمة التوصية وزيادة فهمنا ومعرفتنا بهذا المجال.

## 2-2- الأعمال المتعلقة

انطلاقاً من أهمية أنظمة التوصية وفعاليتها في جذب المستخدمين والمحافظة على ولائهم من خلال تلبية تطلعاتهم، هناك الكثير من الأعمال الملهمة التي يستخدمها ملايين المستخدمين يومياً حول العالم والتي باتت المراجع الأهم في مجالها، ونخص في الذكر التجارب التالية:

### Netflix Prize ومشكلة التوصية

أعلنت نتفليكس عن Netflix Prize في عام 2006 وهي مسابقة في مجال تعلم الآلة وتنقيب البيانات لتوقع تقييم الأفلام. حيث قدمت جائزة مقدارها مليون دولار لأي شخص يحسن دقة النظام الحالي المسمى Cinematch بنسبة 10%. أجريت هذه المسابقة للعثور على طرق جديدة لتحسين التوصيات التي يقومون ب تقديمها لمستخدميهم. ومع ذلك، قاموا

بطرح سؤال افتراضي يكون أسهل في التقييم والقياس: حذر متوسط الخطأ المربع للتقدير المتوقع. بدأ السباق بهدف تحسين قيمة RMSE لدى نتفليكس من (22) إلى 0.8572 أو أقل.

## من DVD إلى البث العالمي

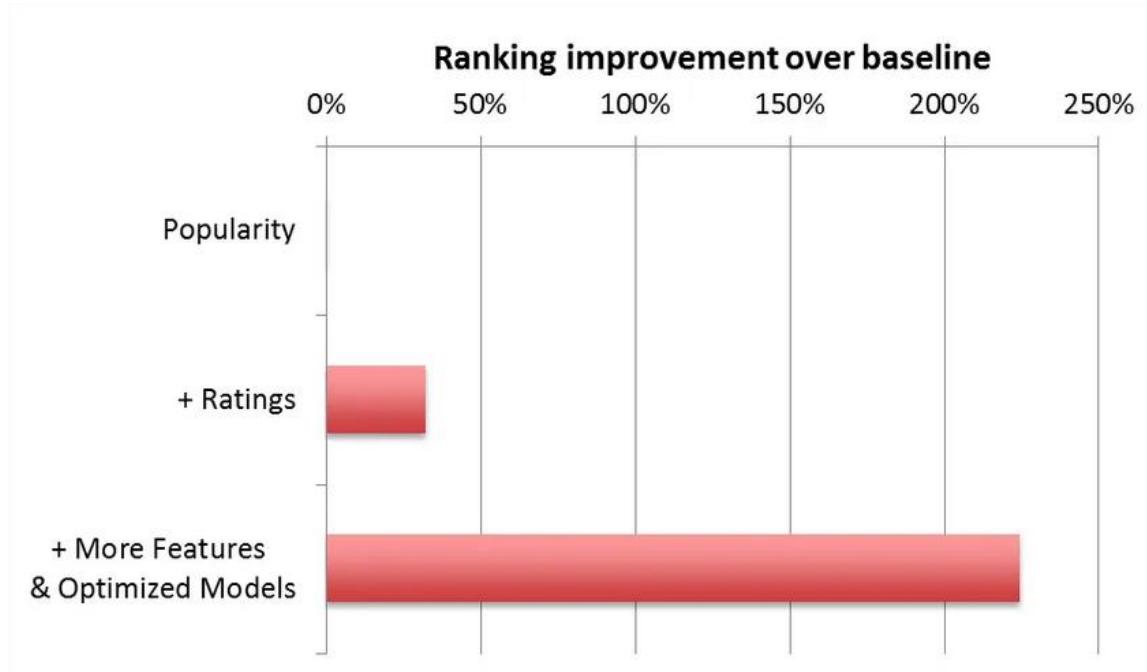
من أسباب تغير تركيز نتفليكس في خوارزميات التوصية هو التغير الكبير الذي طرأ على Netflix كل في السنوات القليلة الماضية. أطلقت Netflix خدمة البث في عام 2007، أي بعد عام واحد من بداية جائزة Netflix. لقد غير البث ليس فقط طريقة تفاعل المستخدمين مع الخدمة، ولكن أيضًا نوع البيانات المتاحة لاستخدامها في خوارزمياتهم. بالنسبة لأفراد الـ DVD كان هدفنا هو مساعدة الأشخاص في ملء قائمة بالعناوين التي سيتلقونها عبر البريد في الأيام والأسابيع القادمة؛ حيث يكون اختيار الفيلم بعيداً في الوقت عن مشاهدته، حيث يقوم الأشخاص باختياره بعناية لأن تبديل أفراد الـ DVD يستغرق أكثر من يوم، ولا يحصلون على أي ردود فعل أثناء المشاهدة. أما بالنسبة لمستخدمي البث، يبحثون عن شيء رائع لمشاهدته الآن؛ حيث يمكنهم تجربة بعض مقاطع الفيديو قبل أن يستقروا على واحد، يمكنهم مشاهدة العديد منها في جلسة واحدة، مما يمكن Netflix من مراقبة إحصائيات المشاهدة مثل ما إذا كان قد تم مشاهدة الفيديو بشكل كامل أو جزئياً.

تغير آخر كبير كان بانتقالنا من موقع واحد إلى مئات الأجهزة. تم الإعلان عن التكامل مع جهاز Roku وXbox في عام 2008، أي بعد عامين من مسابقة Netflix. وبعد عام واحد فقط، وصل البث على Netflix إلى iPhone. اليوم، يتتوفر على مجموعة من الأجهزة التي تمتد من العديد من الأجهزة Android إلى أحدث AppleTV.

قبل عامين، انطلقت Netflix على الصعيدين الوطني والدولي بإطلاق الخدمة في كندا. في عام 2011، قاموا بإضافة 43 دولة وإقليماً لاتينياً إلى القائمة. وفي عام 2012، قاموا بإطلاق الخدمة في المملكة المتحدة وأيرلندا. اليوم، يمتلك Netflix أكثر من 23 مليون مشترك في 47 دولة. استمتع هؤلاء المشتركون بتدفق 2 مليار ساعة من محتوى متتنوع عبر مئات الأجهزة المختلفة في آخر ربع من عام 2011. يقومون يومياً بإضافة 2 مليون فيلم وعرض تلفزيوني إلى قائمة الانتظار ويقومون بإصدار 4 ملايين تقييم.

لقد قاموا بملائمة خوارزميات التخصيص لديهم مع هذا السيناريو الجديد بطريقة تجعل 75% مما يشاهده الناس هو نتیجة لأحد أنواع التوصية. تم ذلك من خلال تحسين تجربة الأعضاء بشكل مستمر، وقد قاسوا مكاسب كبيرة في رضا الأعضاء كلما قاموا بتحسين التخصيص لمستخدميهم.

حققت جائزة نتفليكس (Netflix Prize) تسلیط الضوء على أهمية واستخدام أنظمة التوصية في تطبيقات العالم الحقيقي. قدمت المسابقة العديد من الدروس حول كيفية التعامل مع التوصية، وتعلمت المزيد منذ منحت الجائزة الكبرى في عام 2009. كان تطور التطبيقات الصناعية لأنظمة التوصية يدفعه توفر أنواع مختلفة من بيانات المستخدم ومستوى الاهتمام بهذا المجال في مجتمع البحث. (23)



الشكل 4: تحسين الترتيب عن الحالة الأولية

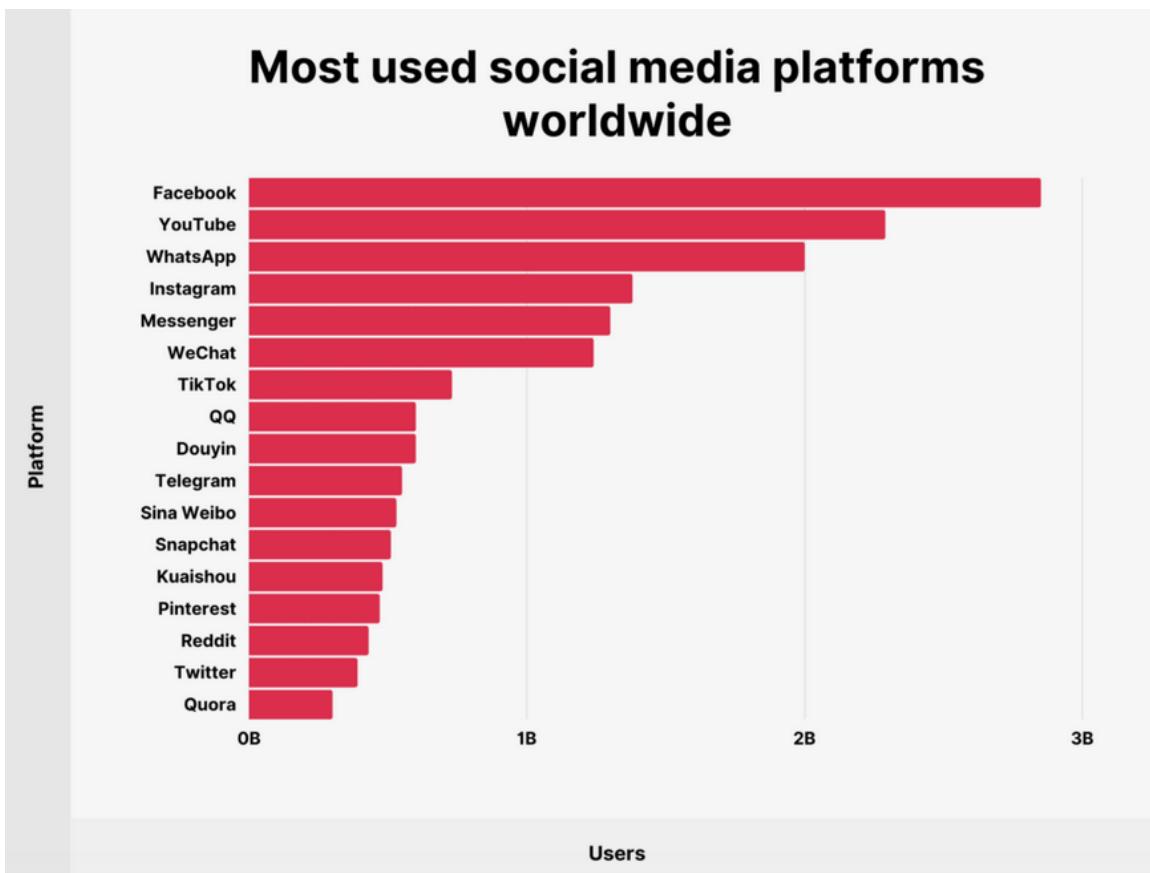
## YouTube

تُعدُّ YouTube منصة فريدة ومثيرة لمشاركة واستكشاف الفيديوهات عبر الإنترنت المصنفة الثانية عالمياً من بين منصات وسائل التواصل الاجتماعي بعد فيسبوك في عام 2020 المرتبة

حسب عدد المستخدمين النشطين. إنها موطن لملايين الفيديوهات التي تغطي كل مجالات الإبداع والترفيه، من مقاطع الفكاهة الطريفة إلى الفيديوهات التعليمية الشيقة وحتى الأحداث العالمية الهامة. يعتبر YouTube مكاناً يجمع بين محتوى المستخدمين والمحتوى المهني، حيث يمكن للأفراد والشركات على حد سواء مشاركة إبداعهم وفنونهم مع العالم أجمع. سواء كنت تبحث عن طرق جديدة للتسلية أو ترغب في اكتساب المعرفة والمهارات، يقدم YouTube مجموعة متنوعة من الفيديوهات التي تلبى تلك الاحتياجات. مع ميزة البحث الفعالة والتصنيفات المتعددة، يمكنك اكتشاف محتوى جديد ومثير يناسب اهتماماتك بسهولة، ومن أهم النقاط التي تساعد في ذلك:

(24)

1. تكامل مع وسائل التواصل الاجتماعي: يمكن لنظام التوصية في يوتيوب أن يستفيد من معلومات من وسائل التواصل الاجتماعي لتحسين دقة التوصيات. يمكن أن تشمل هذه المعلومات اتجاهات الأصدقاء، وتفاعلاتهم، وتفضيلاتهم. وذلك من خلال فهم ما يشاهده ويستمتع به أصدقاؤك.
2. البيانات الضخمة: نظراً لكمية البيانات الكبيرة المستخدمة في نظام التوصية، يعتمد بشكل كبير على معالجة كميات ضخمة من تاريخ المشاهدات، وتفاعلات المستخدم، وتقييماتهم. هذا يساعد على تدريب نماذج التعلم الآلي لتحسين دقة توصيات الفيديوهات بناءً على تفاصيل متقدمة من سلوك المشاهدة
3. تجربة من خلال مختلف الأجهزة: بما أن يوتيوب متاح على مجموعة متنوعة من الأجهزة مثل الهواتف الذكية والأجهزة اللوحية والأجهزة التلفزيونية، يجب أن يتم تكامل نظام التوصية بشكل فعال مع هذه البيئات المتنوعة. مما يعزز راحة المستخدم وقابلية استخدام المنصة
4. التوعية وشرح التوصيات: هدف نظام التوصية إلى جعل المستخدمين على دراية بكيفية تكييفه مع تفضيلاتهم. يُشجع المستخدمون على إعطاء ردود فعل لتحسين التوصيات. هذا يعزز الثقة في النظام ويشجع المستخدمين على المشاركة، مما يؤدي إلى تحسين مستمر في جودة التوصيات.



الشكل 5: انتشار منصات التواصل الاجتماعي

من خلال الدراسات السابقة والأعمال والتجارب الملموسة ومساهمتهم في فهم أهمية أنظمة التوصية وكيف يمكن أن تلعب دوراً حاسماً في جذب المستخدمين والحفاظ على ولائهم من خلال تلبية تطلعاتهم. ومن خلال استخدامنا يومياً لمنصات التواصل الاجتماعي ومنصات الأفلام، حاولنا في هذا المشروع محاكاة هذه التجارب والاستفادة منها.

### 3- الفصل الثالث: القسم العملي

#### 3-1 التحليل

في إطار التطور المستمر لميدان الترفيه الرقمي، تأخذ أنظمة التوصية دوراً بارزاً، ولا سيما في مجال موضع الأفلام، مستفيدة من نجاحات مشروعات متقدمة كجائزة نتفليكس. بالإضافة إلى ذلك، تُظهر النظم المتقدمة لـتوصية المحتوى، المستخدمة يومياً على منصات رائدة مثل يوتيوب، فيسبوك، وتيك توك، قدرة فائقة على إلهام المجال. تلعب هذه الأنظمة دوراً حيوياً في تعزيز تجربة المستخدم من خلال تقديم اقتراحات شخصية و المناسبة للأفلام، مما يسهم بشكل فعال في تعزيز رضا المستخدم وتعاظم مشاركته. يصبح الاعتماد على أنظمة التوصية أمراً أساسياً لمساعدة المستخدمين في التنقل بين العروض السينمائية الواسعة واستكشاف الأفلام المتtagمة مع ذوقهم وفضoliاتهم الشخصية.

مع ذلك، ورغم التبني المتزايد لأنظمة التوصية الفعالة على العديد من المنصات، نجد أنفسنا أمام تحديات وفجوات تظل قائمة في الوقت الحالي. يتمثل أحد هذه التحديات في عدم وجود نظام توصية، أو في فقدان موضع الأفلام لنظام توصية فعال. بالإضافة إلى ذلك، يُشكل تحسين دقة التوصيات تحدياً رئيسياً، بالإضافة إلى ضرورة فهم تفاصيل تفضيلات المستخدمين بشكل أعمق، وتكامل متطلبات الوقت الفعلي لتحديد الاتجاهات والتغييرات في تفضيلات المستخدمين.

يتطلع هذا المشروع إلى محاكاة تلك النجاحات، مع التركيز البارز على تعزيز التطوير المستمر داخل أنظمة التوصية. باستخدام خوارزميات وتقنيات حديثة، نسعى إلى الوصول إلى الدقة والاستجابة في توفير اقتراحات أفلام متغيرة تتtagم مع تفضيلات المستخدمين المتغيرة. يسعى هذا الجهد المشترك إلى تسليط الضوء على أهمية التوجه نحو التطوير المستمر لأنظمة التوصية، من خلال معالجة الفجوات الحالية وتحديد مستقبل اكتشاف المحتوى الشخصي في ميدان الترفيه الرقمي.

#### 3-1-1 الهدف من المشروع

يهدف مشروعنا إلى إنشاء موقع ويب يعرض تفاصيل الأفلام بشكل شامل، مع تطوير نظام توصية فعال. ينبع الهدف من تحسين تجربة المستخدم، حيث يهدف المشروع إلى توفير منصة متميزة لاستعراض الأفلام واكتشاف محتوى جديد بناءً على توصيات دقيقة ومتخصصة.

## 3-1-2- متطلبات المشروع

سنتطرق في تحليل المشروع إلى النقاط الأساسية التالية:

### I. الجمهور المستهدف

يستهدف موقعنا الإلكتروني للأفلام جمهوراً متنوعاً ذو تفضيلات وعادات مشاهدة متنوعة. يشمل الجمهور المستهدف:

**عشاق الأفلام:** الأفراد الذين يشغفون باكتشاف مجموعة واسعة من الأفلام عبر مختلف الأنواع واللغات

**المشاهدين العابرين:** المستخدمون الذين يبحثون عن ترفيه عابر ومهتمون باكتشاف الأفلام الرائجة والشائعة.

**الجماهير ذات التفضيلات النوعية:** الأفراد ذوو التفضيلات لأنواع معينة، مثل عشاق الأكشن، والدراما، والكوميديا، والإثارة، والرومانسية، أو عشاق الوثائقيات.

**الفئات العمرية:** تلبية احتياجات الترفيه لفئات عمرية مختلفة، بما في ذلك المراهقين والشباب البالغين والمشاهدين البالغين.

**عشاق السينما:** عشاق السينما الذين يقدرون الفن السينمائي والأفلام المستقلة والسينما الكلاسيكية.

**المستخدمون الملمون بالเทคโนโลยيا:** الأفراد الذين يشعرون بالراحة في استخدام المنصات الإلكترونية، بما في ذلك أولئك الذين يعرفون خدمات البث واستهلاك المحتوى الرقمي.

**العائلات:** المستخدمون الذين يبحثون عن أفلام مناسبة لجميع الفئات العمرية.

**عشاق الأفلام الدولية:** الجماهير المهتمة باستكشاف أفلام من مختلف الدول والثقافات

## **فضائل المستخدم**

فهم فضائل جمهورنا المستهدف يعد أمراً حاسماً لتقديم تجربة توصية سينمائية شخصية وجذابة. العوامل التي تؤثر في فضائل المستخدم تشمل:

**فضائل النوع:** تحليل شعبية أنواع معينة بين المستخدمين وتصنيف التوصيات وفقاً لذلك.

**أنماط التقييم:** فحص أنماط تقييم المستخدم لفهم الأذواق الفردية وفضائلهم للأفلام المرتفعة التقييم أو الأفلام الخاصة.

**سجل المشاهدة:** استخدام البيانات التاريخية لتوصية الأفلام استناداً إلى عادات المشاهدة السابقة وتفاعلات المستخدم مع المنصة.

من خلال فهم احتياجات وفضائل جمهورنا المستهدف بشكل شامل، يهدف نظام التوصية لدينا إلى تقديم تجربة اكتشاف أفلام شخصية وممتعة بشكل كبير

## **II. متطلبات مجموعة البيانات المستخدمة**

تعتبر كون بيانات الأفلام كبيرة وشاملة أمراً ضرورياً لتحقيق فعالية نظام التوصية، حيث يضمن ذلك توفير مجموعة واسعة وتنوع في الأفلام، مما يسهم في تحسين دقة وفاعلية النظام. بيانات الأفلام الكبيرة تساعد في التنوع في التصنيفات وتوفير تفاصيل شاملة حول عناوين الأفلام، تصنيفاتها، وتقييمات المستخدمين، مما يعزز فهم السياق ويحسن الدقة.

هذه البيانات تتيح تلبية احتياجات مستخدمين متعددين من خلال توفير أفلام بمخالف اللغات والثقافات، مما يسهم في فهم اتجاهات سوق الأفلام وتحليل المحتوى. بالإضافة إلى ذلك، يمكن تحديث البيانات الدوري من تحسين تنبؤات النظام وتوفير توصيات أكثر دقة.

تعتبر حداة البيانات أمراً حيوياً لتحسين كفاءة نظام التوصية. يجب أن تكون البيانات حديثة لتوفير توصيات دقيقة وملهمة، مما يتيح إدراج الأفلام الحديثة ومتابعة تقييمات وآراء المستخدمين، وبالتالي تعزيز تجربة المستخدم وتقديم محتوى متوع ومحدث بانتظام.

### III. متطلبات صفحات الويب

كما ذكرنا في المقدمة من معايير تقيس مدى كفاءة موقع الويب من وجهة نظر المستخدم والمبرمج. وبما أننا موقع الويب خاصتنا هو موقع أفلام وموجه ليستقطب مجموعة متنوعة من المستخدمين. ف لا بد من التأكيد على المتطلبات التالية:

#### متطلبات واجهة المستخدم

1. تصميم جذاب وسهل التفاعل لتعزيز تجربة المستخدم في البحث والاستعراض.
2. إنشاء قائمة تنقل فعالة ومنظمة لتحقيق سهولة التصفح.  
تطوير صفحات تفاصيل الأفلام بمحوى شامل ومعلومات دقيقة وصور ذات دقة عالية.
3. ضمان سهولة البحث وتوفير نتائج دقيقة لضمان الوصول السريع للمحتوى.

#### متطلبات تقنية

1. تصميم متقارب يمكن الوصول إليه بسهولة عبر مختلف الأجهزة.
2. تكامل مع قواعد البيانات لتخزين البيانات المتعلقة بنشاط المستخدم دون التأثير على تجربته.
3. استخدام تقنيات متقدمة لتحميل الصور والمحوى من خلال API بشكل فعال لتعزيز سرعة وأداء الصفحات.

## IV. خطوات تصميم أنظمة التوصية

1. **جمع البيانات:** تتطلب أنظمة التوصيات كميات كبيرة من البيانات لعمل تنبؤات دقيقة. يمكن جمع البيانات من تفاعلات المستخدم مع العناصر، مثل تاريخ سجل البحث، وعدد مرات زيارة المستخدم لفيلم معين من سجل البحث.

2. **المعالجة المسبقة:** قبل أن يمكن تطبيق خوارزميات التوصية، يجب معالجة البيانات مسبقاً. يمكن أن يشمل ذلك تنظيف البيانات وتصفيتها لإزالة الضوضاء والمعلومات غير ذات الصلة. قد تحتاج البيانات أيضاً إلى تحويلها إلى تنسيق مناسب للخوارزمية.

3. **بناء النظام:** بناء نظام توصية هجين يجمع بين نظامي توصية، التصفية التعاونية والتصفية القائمة على المحتوى.

4. **اختبار النظام:** قياس كفاءة النظام من خلال الاختبارات وتحليل توافقه وتقاربه مع احتياجات المستخدم.

5. **قابلية التطوير:** القدرة على استخراج مجموعة البيانات الحالية وتاريخ تفاعلات المستخدم مع الزمن، مما يمكن المطورين مستقبلاً من الاستفادة من تلك البيانات في تطوير نظام توصية أكثر تقدماً ودقةً.

ما يجب مراعاته أيضاً في نظام التوصية:

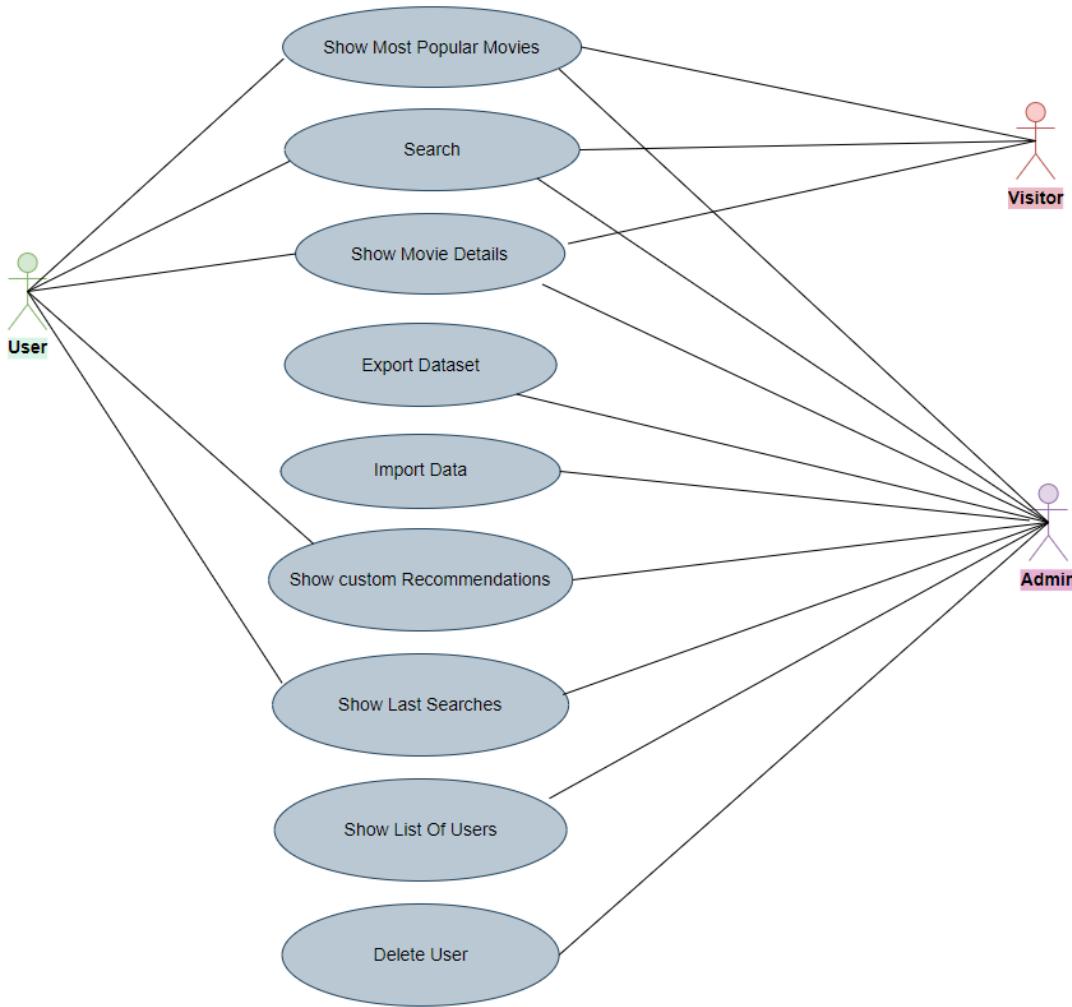
1. **الوصيات في الوقت الفعلي:** زيادة أهمية توفير توصيات فورية بتحديثات في الوقت الحقيقي لتلبية احتياجات المستخدمين.

2. **التنوع:** زيادة أهمية توفير توصيات فورية بتحديثات في الوقت الحقيقي لتلبية احتياجات المستخدمين.

3. **التحسين متعدد الأهداف:** استخدام تقنيات التحسين متعددة الأهداف لتحقيق توازن بين دقة التوصيات والتنوع والحداثة.

### 3-2- التصميم

#### 3-2-1 مخطط حالة الاستخدام Use Case Diagram



في تصميم نظام توصية الأفلام الخاص بنا، لدينا ثلاثة من المستخدمين:

1- الزائر (Visitor):

يُمثل الأفراد الذين يتفاعلون مع النظام دون إنشاء حساب أو تسجيل الدخول.

2- المستخدم المسجل (Registered User):

يُمثل الأفراد الذين قاموا بإنشاء حساب على المنصة وتسجيل الدخول وليس لهم صفة مدير.

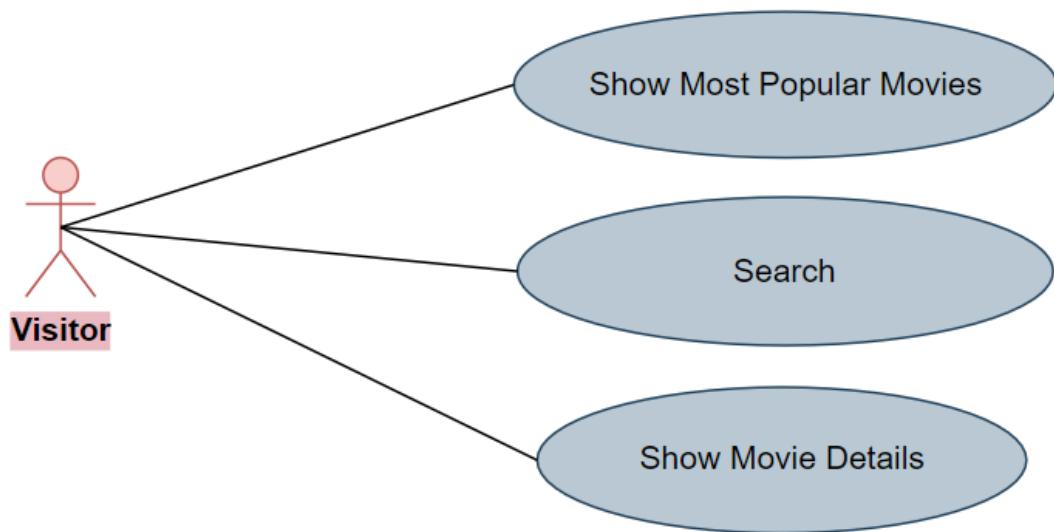
### 3- المدير (Admin):

يمثل الشخص الذي لديه حساب وقام بتسجيل الدخول ويحمل صفة مدير.

تمثل هذه الفصل بين أنواع المستخدمين فرصة لتكيف تجربة المستخدم، مما يوفر تجربة سلسة وشخصية للمستخدمين المسجلين، في حين يضمن واجهة ودية وإعلامية للمستخدمين غير المسجلين. فيما يلي سنقوم بالتفصيل في الوظائف وحالات الاستخدام الخاصة بكل دور مستخدم، ملقية الضوء على تفاصيل نظام توصيات الأفلام الخاص بنا.

حالات الاستخدام بحسب كل مستخدم:

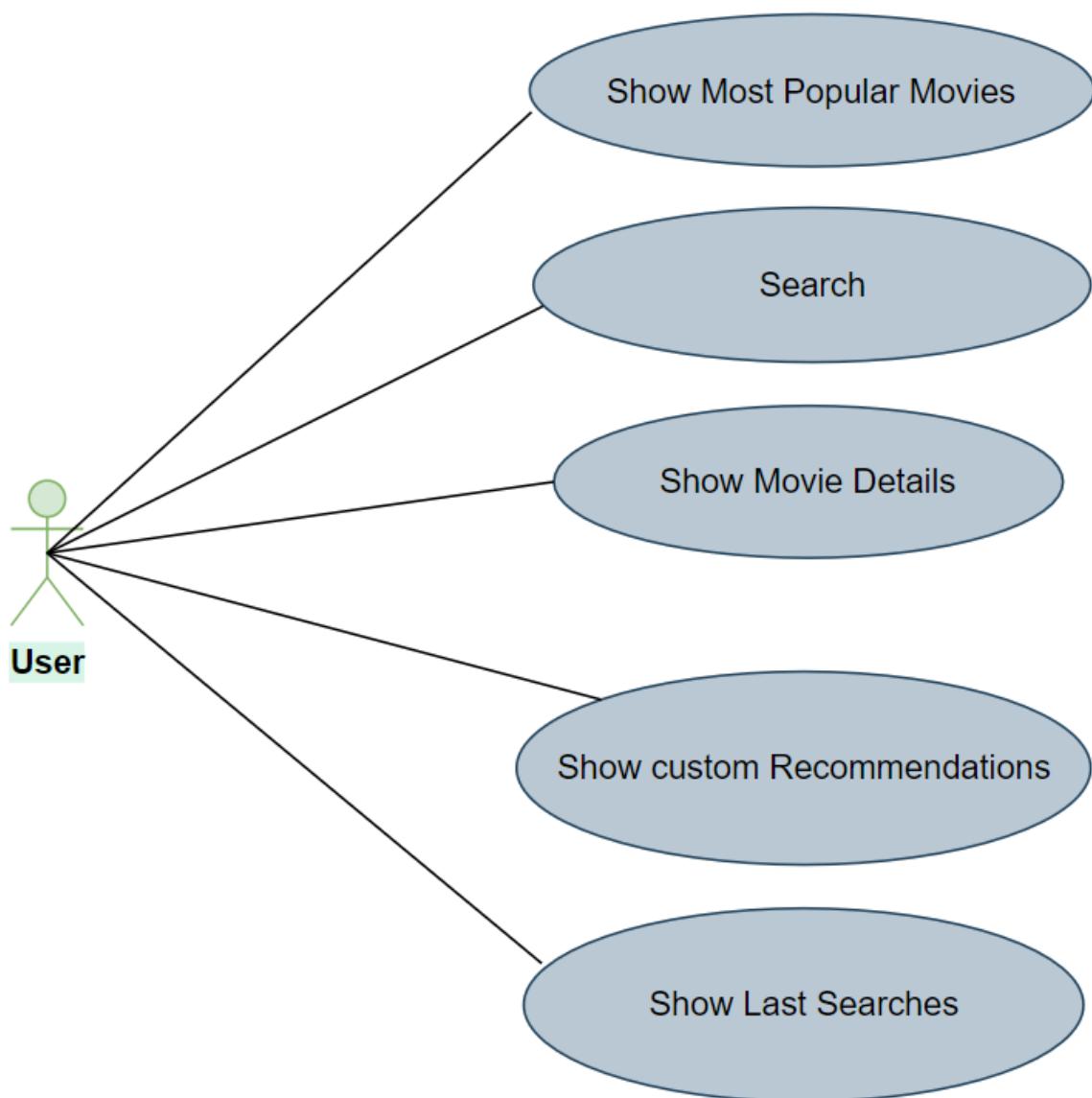
الزائر:



بإمكان الزائر حتى ولو لم يقم بتسجيل الدخول:

- استعراض أكثر الأفلام شهرة.
- البحث عن فيلم معين.
- الاطلاع إلى تفاصيل الأفلام.

المستخدم المسجل:



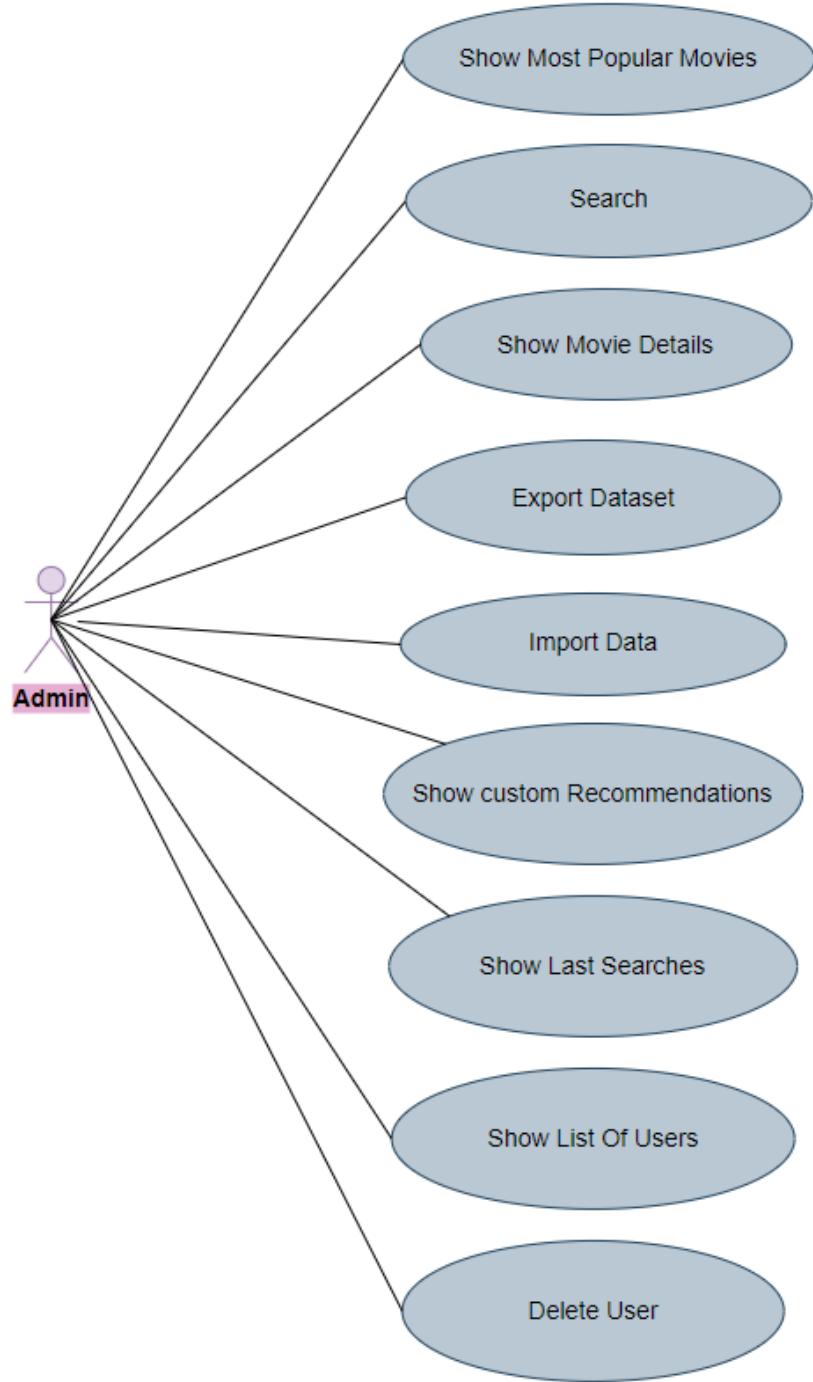
بإمكان المستخدم بعد قيامه بالتسجيل في الموقع والدخول إلى حسابه:

- الاطلاع على أكثر الأفلام شهرةً.
- البحث عن فيلم معين.
- الاطلاع على تفاصيل الأفلام.
- الحصول على توصيات بناءً على الأفلام التي قاموا بالبحث عنها مسبقاً.
- الاطلاع على آخر سجلات البحث التي قاموا بها.

المدير:

مدير الموقع الذي يكون مسجل في الموقع ويحمل صفة مدير بإمكانه القيام بما يلي:

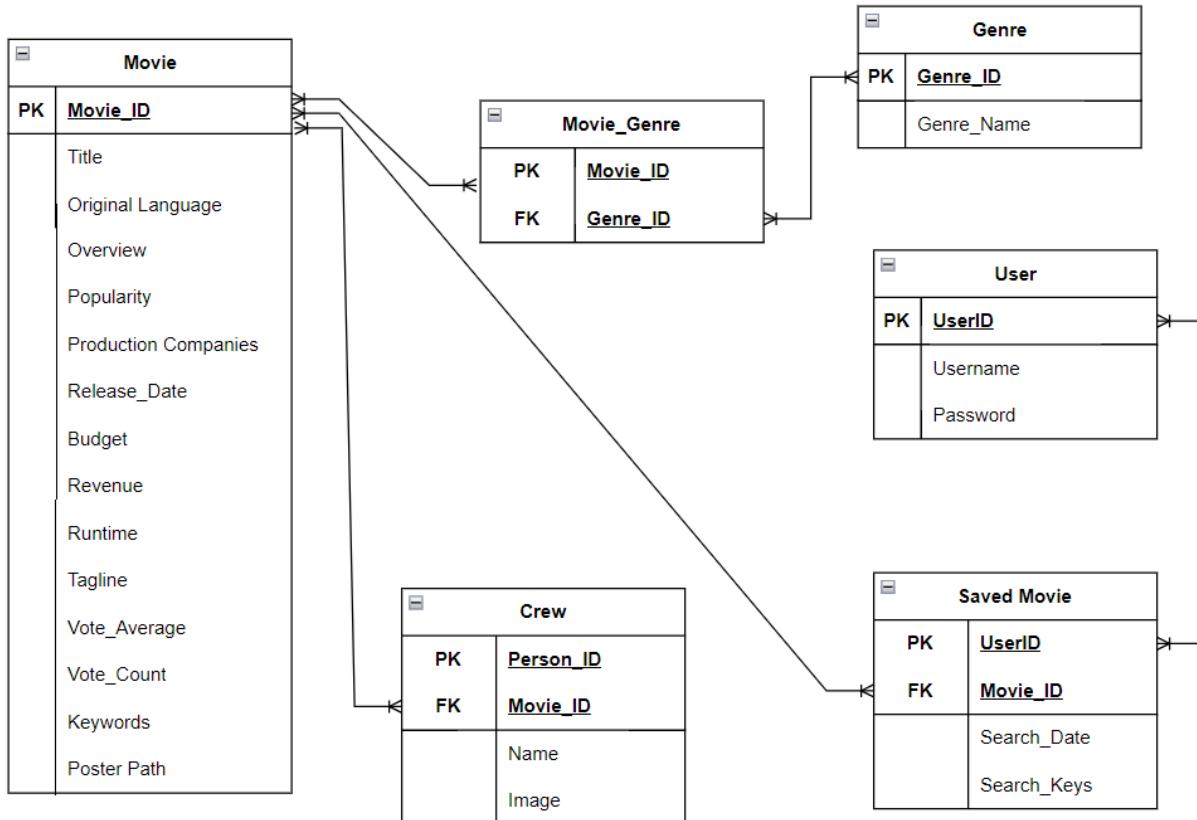
- الاطلاع على أكثر الأفلام شهرةً.
- البحث عن فيلم معين.
- الاطلاع على تفاصيل الأفلام.
- الحصول على توصيات بناءً على الأفلام التي قاموا بالبحث عنها مسبقاً.
- الاطلاع على آخر سجلات البحث التي قاموا بها.
- الاطلاع على المستخدمين غير المدراء المسجلين في الموقع.
- حذف مستخدم.
- تصدير البيانات.
- إدخال مجموعة بيانات إضافية.



تتيح هذه الحالات فهماً شاملًا لكيفية تفاعل المستخدمين مع نظام توصية الأفلام لدينا، وتبهر الجوانب والميزات التي يقدمها النظام لكل نوع من أنواع المستخدمين.

## ERD Diagram -3-2-2

نقوم من خلال مخطط ERD بتوسيع الهيكل العام لقاعدة البيانات وال العلاقات بين الكيانات.



بما أن مخطط الهيكل الكياني (ERD) يعد جزءاً أساسياً من تصميم قاعدة البيانات للموقع، وبالتالي ف هذا المخطط يقوم بتمثيل العلاقات والتفاعلات بين الكيانات المختلفة في النظام. يشمل ERD كل من الكيانات وال العلاقات بينها.

**الكيانات:**

الفيلم Movie: يمثل الفيلم الكيان الرئيسي في الموقع، حيث يحتوي على معلومات مفصلة حول الأفلام مثل عنوان الفيلم، النوع والتقييم.

التصنيف genre: يمثل الأنواع المختلفة للأفلام مثل الأكشن، المغامرة، الكوميدي، الخ.

طاقم العمل Crew: يحتوي على أسماء الممثلين في كل فيلم.

المستخدم User: يحتوي على معلومات حول المستخدمين المسجلين في الموقع بما في ذلك اسم المستخدم وكلمة المرور.

تصنيف الفيلم Movie\_Genre: يحتوي على الفيلم وتصنيفاته، حيث يحتوي على معلومات حول النوع الخاص بكل فيلم.

الأفلام المحفوظة Saved Movie: يمثل سجل الأفلام التي تم حفظها من قبل كل مستخدم بالإضافة إلى عدد مرات تكرار البحث عنها.

وفيما يلي سنقوم باستعراض السمات Attributes الخاصة بكل كيان Entity من المخطط السابق:

### كيان الفيلم :Movie

- معرف رئيسي لكل فيلم Movie\_ID:
- عنوان الفيلم Title:
- اللغة الأصلية للفيلم Original\_Language:
- وصف أو ملخص عن الفيلم Overview:
- درجة شهرة الفيلم Popularity:
- الشركات المشاركة في إنتاج الفيلم Production\_Companies:
- تاريخ إصدار الفيلم Release\_Date:
- الميزانية المخصصة للفيلم Budget:
- إيرادات الفيلم Revenue:
- مدة الفيلم Runtime:
- الحالة الحالية للفيلم (مثل: تم الإصدار، قيد الإنتاج) Status:
- شعار الفيلم أو الجمل الدعائية المرتبطة بالفيلم Tagline:
- متوسط التصويت أو التقييم للفيلم Vote\_Average:
- عدد التصويتات أو التقييمات التي حصل عليها الفيلم Vote\_Count:
- الأشخاص أو الجهات المساهمة في الفيلم Credits:

- **Keywords:** الكلمات الرئيسية المرتبطة بالفيلم
- **Poster\_Path:** مسار ملف الصورة الرئيسية للفيلم

**كيان التصنيف :Genre**

- **Genre\_ID:** معرف رئيسي لكل نوع أو تصنيف
- **Name:** اسم النوع أو التصنيف

**كيان طاقم العمل :Crew**

- **Person\_ID:** معرف رئيسي لكل ممثل
- **Movie\_ID:** معرف ثانوي يشير إلى الفيلم الذي شارك فيه
- **Name:** اسم الممثل
- **Image:** صورة الممثل

**كيان المستخدم :User**

- **User\_ID:** معرف رئيسي لكل مستخدم
- **Username:** اسم المستخدم
- **Password:** كلمة المرور

**كيان تصنيف الفيلم :Movie\_Genre**

- **Movie\_ID:** معرف ثانوي يشير إلى الفيلم
- **Genre\_ID:** معرف ثانوي يشير إلى النوع أو التصنيف

**كيان الأفلام المحفوظة :Saved Movie**

- **User\_ID:** معرف ثانوي يشير إلى المستخدم المسجل الذي قام بالبحث ومشاهدة فيلم ما
- **Movie\_ID:** معرف ثانوي يشير إلى الفيلم الذي قام المستخدم بالبحث عنه واستعراض تفاصيله
- **Search\_Date:** تاريخ قيام المستخدم بهذا البحث

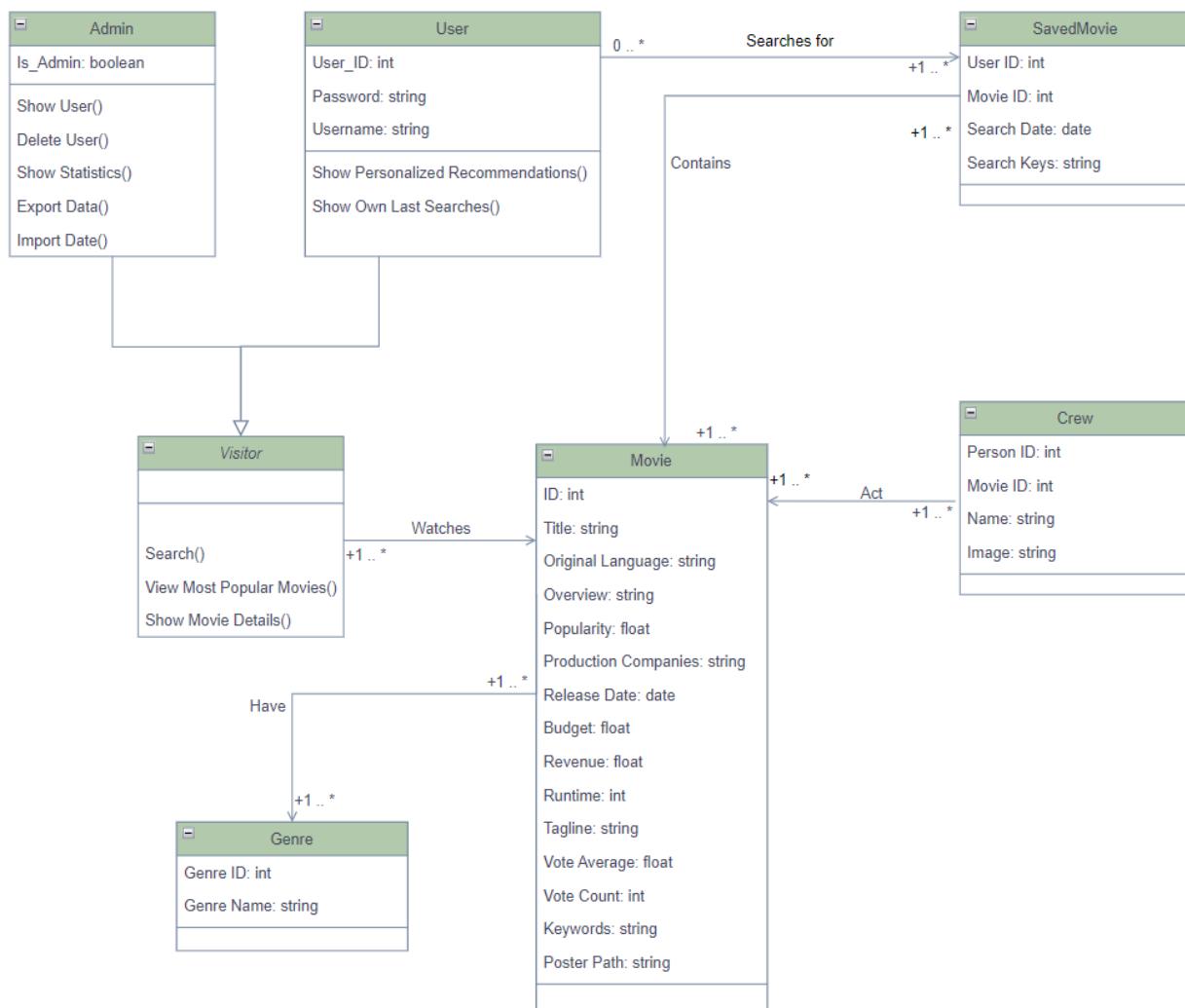
- Search\_Keys: المدخلات التي بحث المستخدم من خلالها

كما يمكن تلخيص العلاقات ما بين الكيانات السابقة والتي هي متعدد لمتعدد في مخطط ERD الخاص بموقع الأفلام الخاص بنا كما يلي:

- ❖ يمكن أن يكون لكل فيلم أكثر من تصنيف أو نوع واحد. على سبيل المثال، يمكن لفيلم مثل Home Alone أن يحمل التصنيفات التالية: كوميدي، عائلي، مغامرة. ومن الممكن أن يندرج تحت أكثر من تصنيف وهذا يختلف على حسب معيار التصنيف الذي يتبعه كل موقع وحسب طبيعة الفيلم ووجهة النظر التي يندرج تحتها التصنيف.
- ❖ كما يمكن لنفس التصنيف أن يطلق على فيلم أو أكثر. على سبيل المثال، تصنيف الفيلم المسما بالرعب يمكن أن يندرج تحته الكثير من الأفلام.
- ❖ بما أن الفيلم هو نتاج جهد طاقم عمل كبير، وبالتالي فإن العلاقة بين كيان طاقم العمل وكيان الفيلم هو متعدد لمتعدد لأنه من الممكن للممثل أن يشارك في أكثر من فيلم كما يمكن للفيلم نفسه أن يضم أكثر من ممثل.
- ❖ يمكن أن يبحث المستخدم ويشاهد الكثير من الأفلام، ويمكن أن يتم البحث عن نفس الفيلم من قبل الكثير من المستخدمين.
- ❖ وأخيراً، يمكن للفيلم الواحد أن يتم البحث عنه من قبل عدد كبير جداً من المستخدمين.

### 3-2-3- مخطط الصف Class Diagram

يُقدم مخطط الصف التالي نظرة شاملة حول تنظيم البيانات في الموقع. يُظهر هذا المخطط كيفية تمثيل المستخدمين والأفلام وسجلات بحث المستخدمين في قاعدة البيانات، مما يسهل عمليات الاستعلام والتفاعل بين المكونات المختلفة للمشروع. تحتوي الجداول الموجودة في هذا المخطط على المعلومات الرئيسية التي تُستخدم في توفير توصيات دقيقة وفعالة للمستخدمين.



### **3-3- التنجيز**

سنستعرض في هذا الفصل الأدوات البرمجية والمكاتب والمصادر المفتوحة التي استخدمتها في إنشاء الموقع. بالإضافة إلى بناء نظام التوصية.

#### **3-3-1- الأدوات البرمجية المستخدمة**

- MySQL Database .1
- Visual Studio Code Version 1.85.1 .2
- Python3 Programming Language .3
- Django Version 4.2.6 Framework .4
- XAMPP .5

#### **قاعدة البيانات MySQL**

تعتبر قواعد البيانات حجر الأساس في تصميم أي نظام يعتمد على تخزين واسترجاع البيانات بكفاءة. في موقع الأفلام الخاص بنا، تم اختيار قاعدة البيانات MySQL لدورها الرئيسي في تحقيق أهداف النظام.

MySQL هو نظام إدارة قواعد البيانات الذي يستخدم لتخزين واسترجاع البيانات. يعتبر MySQL أحد أنظمة إدارة قواعد البيانات مفتوحة المصدر الأكثر شيوعاً في العالم.

SQL (Structured Query Language) هو لغة برمجة تستخدم لإدارة واستعلام قواعد البيانات. يستخدم MySQL لغة SQL للتفاعل مع البيانات، سواء كان ذلك لإدراج بيانات جديدة، تحديث البيانات الحالية، أو استرجاع المعلومات.

#### **مميزات MySQL**

- موثوقية الأداء: توفر أداءً فائقاً وموثوقية عالية، وهي مصممة لتنفيذ العمليات بفعالية حتى في البيئات ذات الأعباء الكبيرة. هذا يضمن استجابة سريعة للمستخدمين وتشغيل سلس لجميع العمليات.

- توفير الوقت والجهد: تتيح استخدام العديد من الأدوات والميزات التي تسهل عمليات التطوير وإدارة قاعدة البيانات. هذا يوفر الوقت والجهد ويسمح بالتركيز على تطوير ميزات النظام بدلاً من التفاصيل التقنية.
- قوة الاستعلام: يدعم استعلامات قوية ومتقدمة، مما يسمح للنظام بتحقيق تحليل بيانات دقيق وفعال. هذا يعزز قدرة المشروع على تقديم توصيات دقيقة وتجربة مستخدم فعالة.
- سهولة الاستخدام: توفر واجهة سهلة الاستخدام وتوثيق شامل.

تم اختيار MySQL بناءً على توافقها مع متطلبات الموقع الذي سنقوم ببنائه. فهي تدعم بنية البيانات المعقدة وعلاقات المتعدد لمتعدد بشكل أمثل، مما يجعلها مناسبة لتخزين واسترجاع بيانات الأفلام والمستخدمين بشكل متكامل.

## Visual Studio Code Version 1.85.1

هو محرر نصوص خفيف الوزن وقوى مفتوح المصدر تم تطويره بواسطة Microsoft يتميز بالسرعة والقابلية للتخصيص ودعم لعدة لغات برمجة. يستخدم على نطاق واسع كبيئة تطوير متكاملة لمطوري البرمجيات. كما يتميز بالخصائص التالية:

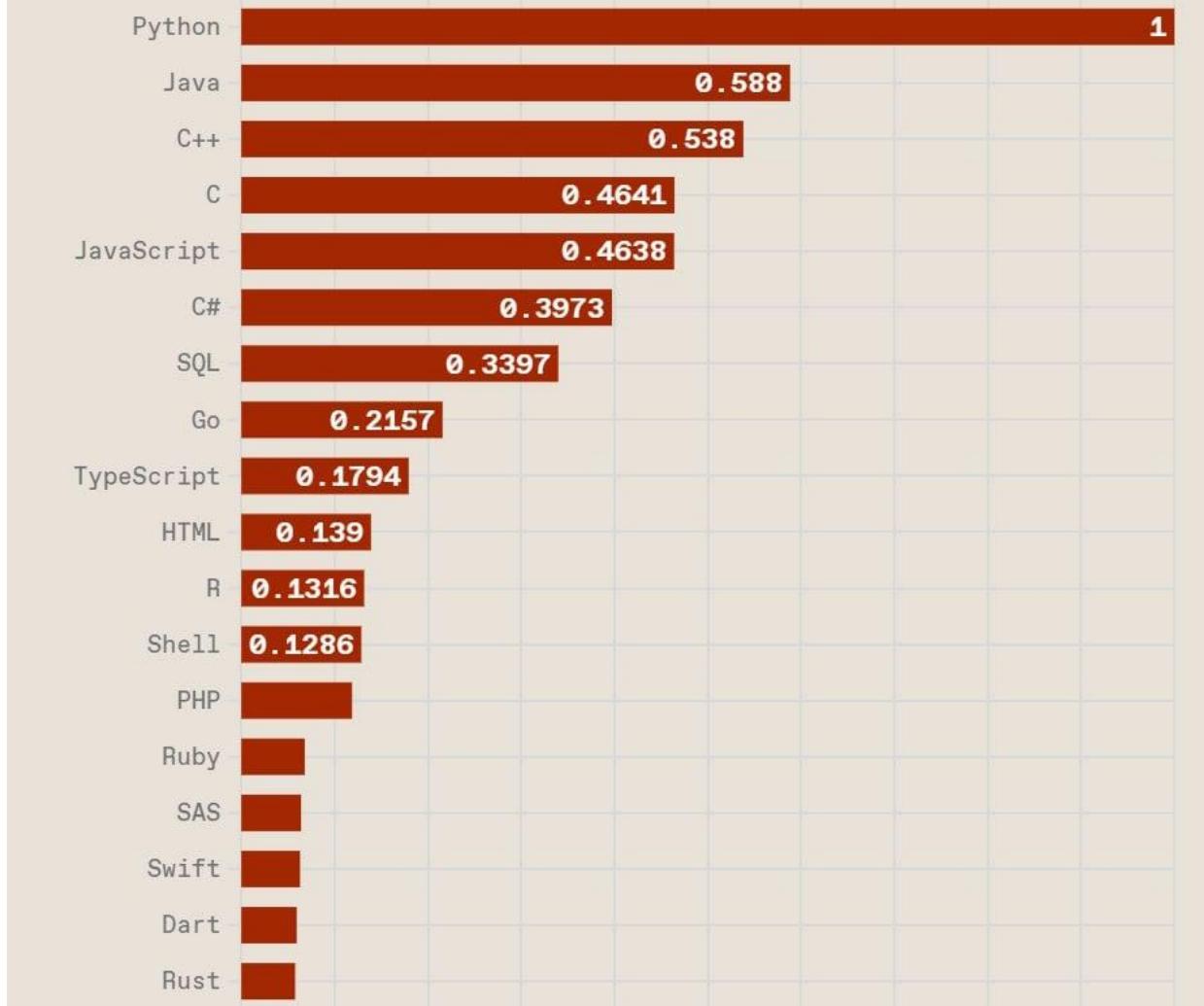
- تحرير متقدم: يتميز VS Code بتحرير متقدم يدعم الكثير من لغات البرمجة، مع تلوين للكود البرمجي وتحديد السطر الحالي، مما يسهل فهم وتحرير الكود.
- التكامل: يدعم Visual Studio Code العديد من اللغات البرمجية الشهيرة، مثل Python، Django، CSS، HTML، JavaScript وغيرها. كما يدعم ملحقات لتحقيق التكامل مع مكتبات وأطر أخرى.
- إدارة المشروع: يوفر واجهة بسيطة وفعالة لإدارة هيكل الملفات والمشاريع. يمكنك التنقل بسهولة بين الملفات والمجلدات والبحث عن الملفات.
- امتدادات وإضافات: يمكن توسيع قدرات VS Code باستخدام مجموعة متنوعة من الامتدادات والإضافات المتاحة. وهذه أحد أهم الميزات التي نحتاجها في مشروعنا في بناء الموقع وفي بناء نظام التوصية.
- التشغيل وفحص الكود: يمكن تشغيل واختبار التطبيقات مباشرةً من البيئة باستخدام أدوات التكامل مع البرمجيات الخارجية. يتيح ذلك تسريع عمليات التطوير والاختبار.

- خفيف الوزن: يعمل ببساطة على معظم الأنظمة دون إهدار موارد النظام.
- سهولة الاستخدام: يتميز التصميم البسيط والديهي لواجهة المستخدم بـ VS Code بترتيب منطقي للأدوات والخيارات. يُسهل تنظيم القوائم والشريط الأدوات الوصول إلى الوظائف المطلوبة بشكل سلس.
- التكامل مع الأدوات الخارجية: يتيح تكامل سلس وفعال مع الأدوات الخارجية مثل Git والمترجمات وغيرها.

## لغة البرمجة Python3

هي لغة برمجة متعددة الاستخدامات ومعتمدة على نطاق واسع تلعب دوراً حيوياً في تطوير أنظمة التوصية، حيث تقدم بيئة غنية من المكتبات والأطر المصممة خصيصاً لعلوم البيانات والتعلم الآلي. بالتنسيق مع بيئة تطوير Python Visual Studio Code (VS Code) ، توفر Python منصة سلسة وفعالة لبناء خوارزميات التوصية. لغتها المفروعة، وموارنتها، والدعم الواسع للمجتمع تسهم في بروزها في مشاريع أنظمة التوصية. باستخدام مكتبات مثل Scikit-Learn و Pandas و TensorFlow، تتيح Python التطوير السريع والتجارب باستراتيجيات متنوعة للتوصية. تعزز التكاملية بين Python و VS Code تجربة التطوير، مما يوفر بيئة قوية للبرمجة والتصحيح والتعاون. في ظل استمرار تطوير أنظمة التوصية، يظل دور Python حاسماً، حيث يمكن للمطورين إنشاء حلول فعالة وقابلة للتوضيع.

# Top Programming Languages 2023

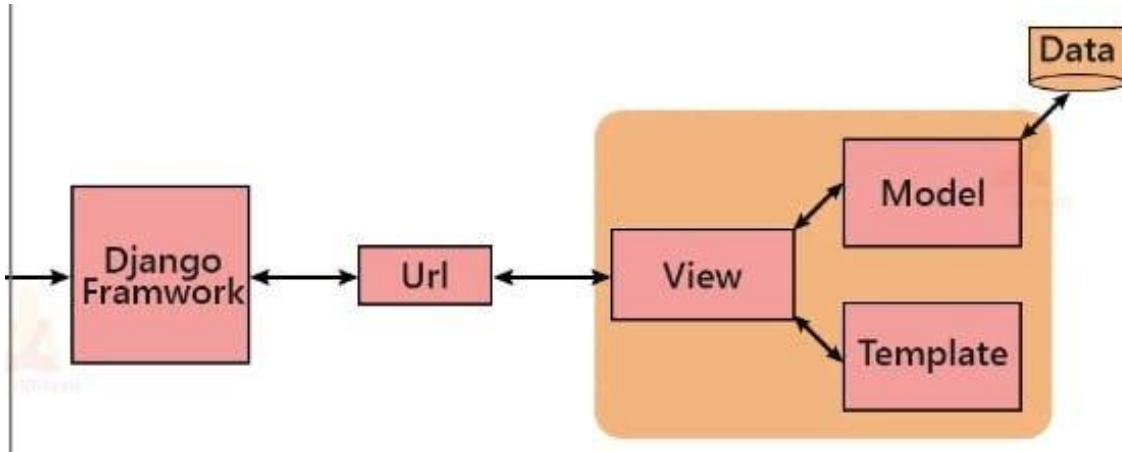


الشكل 6: أشهر لغات البرمجة 2023 حسب IEEE Spectrum

## إطار عمل Django

هو إطار عمل مفتوح المصدر يتبع بنية أساسية تعرف بـ "Model-View-Template". يتيح Django لنا بناء تطبيقات الويب بسرعة وكفاءة. يتبنى هذا الإطار تقنيات تطوير حديثة ويوفر بيئه تطوير مستدامة وقوية. يتمثل تفصيل بنية Django بما يلي:

- **Model**: يستخدم لتمثيل بنية البيانات وقاعدة البيانات. يسهل Django تحديد الـ Models والعلاقات فيما بينها مما يبسط إدارة قاعدة البيانات وسهولة الوصول إليها.
- **View**: تحتوي على المعالجة الرئيسية لما سيتم عرضه لاحقاً للمستخدم من خلال تحديد طريقة استرجاع البيانات من الـ Models وتحديدها للعرض بواسطة الـ Templates.
- **Template**: يستخدم لتحديد وتنسيق هيكلة صفحات الويب بحيث تتيح تخصيص القوالب لجعل واجهة المستخدم جذابة وسهلة الاستخدام.



الشكل 7: بنية *Django*

كما يوفر ميزات قوية لتسهيل عملية التطوير، مثل إدارة قاعدة البيانات، الأمان، قابلية التوسيع. بفضل هيكله الجيد وانتشاره الواسع، يُعد Django خياراً ممتازاً لتطوير تطبيقات الويب المعقدة، بما في ذلك نظم التوصية.

## XAMPP

هو حزمة برمجيات مفتوحة المصدر وقابلة للتوزيع بحرية تستخدم لإعداد وتشغيل بيئة خوادم ويب على أنظمة التشغيل Windows و Linux و macOS. اسم XAMPP يشير إلى المكونات الرئيسية التي تحتوي عليها:

**X (أنظمة التشغيل المختلفة):** تشير إلى أن XAMPP متوفرة للعديد من أنظمة التشغيل، بما في ذلك Windows و Linux و macOS.

**(Apache) A:** يتضمن XAMPP خادم الويب Apache ، الذي يعد واحداً من أشهر خوادم الويب المفتوحة المصدر.

**(MySQL) M:** يحتوي XAMPP على نظام إدارة قواعد البيانات MySQL ، الذي يتيح لك تخزين واسترجاع البيانات.

**(PHP) P:** يأتي XAMPP مزوداً بلغة البرمجة PHP، وهي لغة برمجة قوية وشائعة تستخدم لتطوير تطبيقات الويب الديناميكية.

**(Perl) P:** يتضمن XAMPP Perl، وهي لغة برمجة مستخدمة بشكل شائع في تطوير البرمجيات.

تجعل من السهل إعداد بيئة تطوير محلية لتطوير واختبار تطبيقات الويب بشكل محلي. يسهل XAMPP أيضاً استخدامه لتطبيقات ومشاريع مختلفة، بما أنه يجمع بين مكونات الخادم الأساسية في حزمة واحدة.

استخدمنا XAMPP بالتزامن مع بايثون و Django بسبب الفوائد التي يقدمها في تطوير تطبيقات Django

**تشغيل خوادم متعددة:** يوفر بيئة تشغيل شاملة تشمل خوادم متعددة مثل Apache و MySQL و PostgreSQL. باستخدامها، يمكننا تشغيل تطبيق Django الخاص بي بشكل كامل وفعال.

**دعم لغات برمجة متعددة:** لا يقتصر على دعم لغة معينة، مما يتيح لنا استخدام بايثون و Django بجانب بيئة تشغيل مختلفة.

**قواعد البيانات:** يسهل استخدام XAMPP في إدارة قواعد البيانات، حيث يمكننا تشغيل خادم MySQL بسهولة والتفاعل معه من خلال Django.

**تطوير محلي:** باستخدام XAMPP، يمكننا تنفيذ واختبار تطبيق Django بيناً محلياً، وهو أمر مهم في مراحل التطوير.

**سهولة التثبيت والإعداد:** يوفر XAMPP عملية تثبيت وإعداد بسيطة، مما يساعدنا في تجنب التعقيدات أثناء إعداد البيئة.

**التوافق مع Django:** يمكننا دمج XAMPP بسهولة مع Django واستفادة من قوة الإطار العمل لتطوير تطبيقات الويب بشكل فعال.

### 3-3-2- مراحل العمل على نظام التوصية

#### 3-3-1 جمع البيانات

نظراً لخصوصية وتكلفة الوصول إلى البيانات المطلوبة من خلال معظم الموقع، تم الاعتماد على البيانات الخاصة بموقع TMDB مفتوحة المصدر Kaggle. وذلك لأن موقع TMDB يعتبر موقعًا مفتوح المصدر، حيث تتيح هذه البيانات الوصول إلى معلومات حول الأفلام، الممثلين، وطاقم العمل بشكل شامل.

تم الاعتماد على بيانات ما يقارب 700,000 فيلم بما فيها من عناوين، تواريخ اصدار، شركات الإنتاج، تصنيفات وتقييمات المستخدمين وما إلى ذلك.

بالإضافة إلى الاعتماد على API المجانية الخاصة بموقع TMDB لجلب الصور والممثلين.

### 3-3-2-2 إعداد البيانات

لإنجاز هذه المرحلة تم الاعتماد على مكتبة Scikit-learn، وهي مكتبة مفتوحة المصدر، تدعم خوارزميات تعلم الآلة. تحوي حزم python العلمية مثل:

1. NumPy: Base n-dimensional array package
2. Matplotlib: Comprehensive 2D/3D plotting
3. Pandas: Data structures and analysis
4. Sklearn.feature\_extraction: Extract features from text and images.
5. Sklearn.metrics: Compute score functions, performance metrics, pairwise metrics and distance computations.

لإعداد البيانات قبل البدء بالعمل نقوم بما يلي:



الشكل 8: مراحل إعداد البيانات والتدريب الأولى

قبل استخدام مجموعة البيانات يجب القيام بالتحضير للوصول إلى بيانات عالية الجودة، لذلك قمنا بتنظيف مجموعة البيانات المستخدمة لتصبح جاهزة في مراحل تدريب نظام التوصية لاحقاً. باستخدام إمكانيات مكتبة pandas القوية في لغة البرمجة بايثون.

```
import pandas as pd

file_path = "..\\lMoviesWebsite\\movies_old.csv"
df = pd.read_csv(file_path)

df_cleaned = df.drop_duplicates(subset=['id'])
df_cleaned = df_cleaned[df_cleaned['poster_path'].notna() & (df_cleaned['poster_path'] != '')]

df_cleaned.to_csv('cleaned_data.csv', index=False)
```

## تحميل وتنظيف البيانات

1- إزالة السجلات المكررة، التي تم التعرف عليها بواسطة قيمة المعرف الأساسي للفيلم ID، بشكل منهجي، مما ضمن سلامة وتجنب التكرار ضمن مجموعة البيانات.

2- إزالة السجلات التي تحتوي على قيم فارغة فيما يتعلق بمسار صورة الفيلم Poster Path. هذا الإجراء له أثر إيجابي في القضاء على التكرارات، وتعزيز مجموعة البيانات بالتخلص من الحالات التي تحتوي على معلومات غير مكتملة.

3- بالإضافة إلى ما سبق، قمنا بتحليل بسيط لتصنيفات الأفلام في مجموعة البيانات لفهم توزيع التصنيفات.

```
df['genres'].fillna(' ', inplace=True)

all_genres = ' '.join(df['genres'])

all_genres_list = all_genres.replace('-', ' ').split()

unique_genres = set(all_genres_list)

print(unique_genres)
```

## معالجة البيانات واستخراج الميزات والتوجيه

دمج ميزات الأفلام Features الأساسية وهي ملخص الفيلم، التصنيفات، الممثلين، الكلمات الرئيسية واللغة الأصلية للفيلم (Overview, Genres, Credits, Keywords, Original Language) في قيمة جديدة تسمى الوسوم Tags.

```

df.genres = df.genres.apply(lambda x: " ".join(x.split("-")))
df.keywords = df.keywords.apply(lambda x: " ".join(x.split("-")))
df.credits = df.credits.apply(lambda x: " ".join(x.replace(" ", "").split("-")[:5]))

df["tags"] = df.overview + " " + df.genres + " " + df.credits + " " + df.keywords + " " + df.original_language

new_df = df[["id", "title", "tags", 'poster_path']]

new_df.tags = new_df.tags.apply(lambda x:x.lower())

new_df.head()

```

تنظيف ومعالجة البيانات من خلال التجذير Stemming (وهي عملية إعادة الكلمات إلى جذرها الأساسي مما يسهل فهم السياق وبالتالي تحسين قدرة قياس التشابه بينها. هذا يفيد في تحليل البيانات وفهم الأنماط اللغوية) وتحويلها إلى صيغة صغيرة Lowercase مما يساعد في توحيد البيانات، حيث يتم التعامل بشكل موحد مع الكلمات بغض النظر عن حالة الأحرف (كبيرة أو صغيرة).

تساهم هذه العمليات في تحسين أداء النماذج، حيث يتم تقديم البيانات بشكل أفضل وتحسين قدرتها على فهم السياق وال العلاقات بين الكلمات. بالإضافة إلى تقليل التعقيد Dimensionality Reduction من خلال تقليل الأبعاد بين البيانات مما يساعده في تقليل التعقيد و زمن التحليل. وأخيراً، تسهيل عمليات البحث التي تعزز تجربة المستخدم وتزيد من دقة النتائج.

```

# ### Stemming

ps = PorterStemmer()

def stem(text):
    y = []
    for i in text.split():
        y.append(ps.stem(i))

    return " ".join(y)

new_df["tags"] = new_df["tags"].apply(stem)

```

## عملية الـ Vectorization

وهي عملية تحويل البيانات النصية إلى متجهات رقمية بحيث يمكن استخدامها كمدخلات Input لنموذج نظام التوصية. وهنا قمنا باستخدام CountVectorizer من مكتبة CountVectorizer التي تقوم بتحويل البيانات النصية إلى متجهات رقمية حيث يستخدم كل كلمة كميزة ويتم تحديد قيمة الوزن بناءً على عدد مرات ظهور الكلمة.

تحويل سلسلة النصوص في العمود "tags" - مع إزالة الكلمات الشائعة باللغة الإنكليزية (نظراً لاستخدام مجموعة بيانات باللغة الإنكليزية) وبعد أقصى للسمات المستخلصة هو 2000 سمة - إلى متجهات رقمية باستخدام NumPy و الحصول على مصفوفة fit\_transformNumPy ثنائية الأبعاد.

ومن ثم، استخراج أسماء السمات (الكلمات) المستخدمة في المصفوفة من 80 إلى 84 سمة.

```
# ### Text Vectorization
#
cv = CountVectorizer(stop_words="english",max_features=2000)

vectors = cv.fit_transform(new_df[ "tags" ]).toarray()

cv.get_feature_names_out()[80:85]
```

### 3-2-3- نظام التوصية وتدريب النموذج

قمنا باعتماد بناء نموذجي توصية، الأول نظام توصية قائم على المحتوى والآخر نظام توصية هجين يجمع بين نظام توصية قائم على المحتوى ونظام توصية قائم على التصفية التعاونية.

### تدريب نموذج تشابه جيب التمام Cosine Similarity

تشابه جيب التمام **Cosine Similarity**: مقياس رياضي يستخدم لقياس التشابه بين عناصر معينة استناداً إلى سماتها. في مجال نظام التوصية، ترکز التصفية القائمة على المحتوى على خصائص العناصر وتقضيلات المستخدمين لتقديم توصيات ذات طابع شخصي.

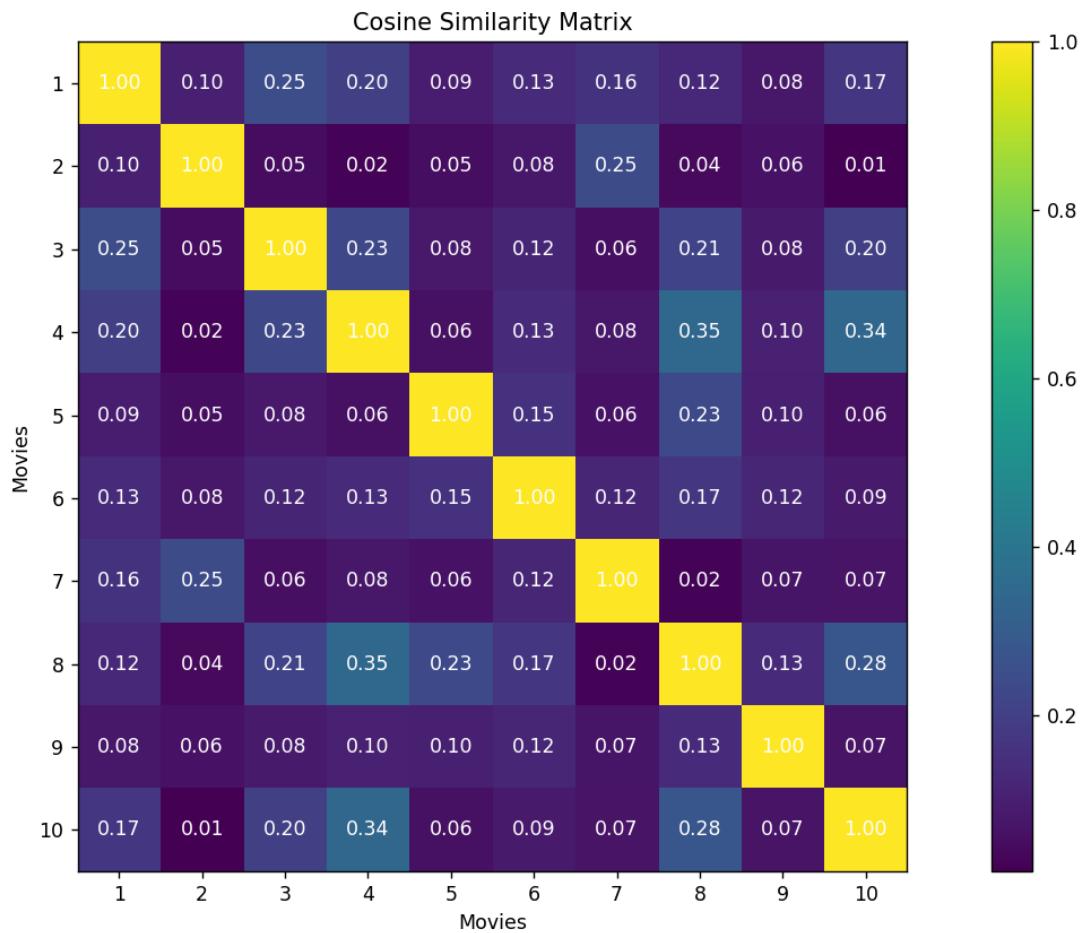
بعدما قمنا بتمثيل العناصر كمتجهات رقمية، فإن تشابه جيب التمام بين عنصرين (متجهين) يتم حسابه من خلال تجيب الزاوية بينهما.

$$\text{Cosine Similarity } (A, B) = (A \cdot B) / (\|A\| * \|B\|)$$

حيث أن:  $(A \cdot B)$  تمثل الضرب النقطي لمتجهي  $A$  و  $B$

$\|A\|$  و  $\|B\|$  يمثلان قيم الطول الإقليدي لكل متجه منهما

**درجات التشابه:** تتراوح درجة التشابه بزاوية الجيب التمام الناتجة في حالة التصفية القائمة على المحتوى من 0 (غير متشابهة تماماً) و 1 (متشابهة تماماً). كما تشير درجة التشابه بزاوية جيب التمام الأعلى إلى وجود تشابه أكبر بين العناصر. كما هو موضح في مصفوفة التشابه التالية في الشكل (9)



الشكل 9: مصفوفة التشابه لعينة من البيانات

توفر Python البرمجية `Sklearn.Metrics.pairwise` دالة قياس تشابه `Cosine_Similarity` .  
وباستخدام حزمة `Pickle` نقوم بتخزين النتيجة في ملف بلاحقة '`pkl`'، للاحتفاظ بالبيانات التي تم استخدامها وحسبها من خلال تدريب نموذج `Cosine_Similarity` على مراحل التوصية اللاحقة مما يضمن لنا تخفيض زمن المعالجة الذي سيترتب على إعادة تدريبيه.

```
similarity = cosine_similarity(vectors)

similarity.shape

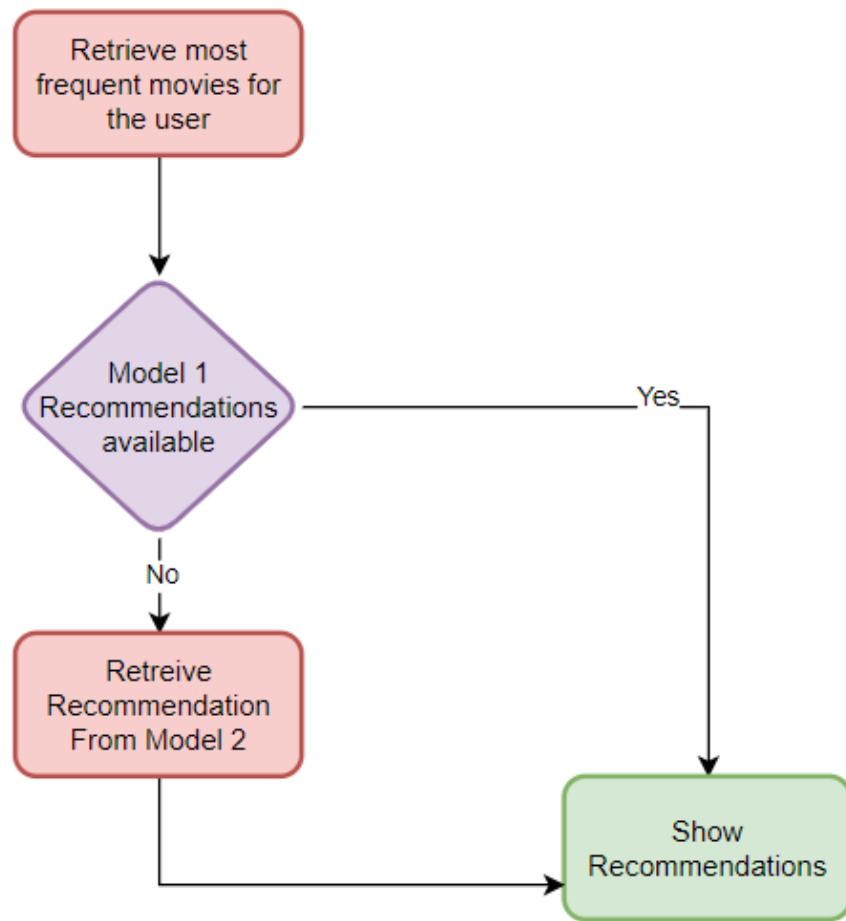
with open('model_data1.pkl', 'wb') as file:
    pickle.dump((new_df, similarity), file)
```

## نهج التوصية الهجين المستخدم

عندما يقوم المستخدم بالدخول إلى أي صفحة ويب من الصفحات التي تقوم بعرض التوصيات في الموقع فإن دالة العرض تقوم بالبحث عن التوصيات الأكثر تشابهاً مع فيلم معين أو مع مجموعة أفلام قام بالبحث عنها مسبقاً من خلال نموذج التوصية الأول.

إذا كانت التوصيات التي تم جلتها بواسطة النموذج الأول، يقوم بعرضها مباشرة للمستخدم.

ولكن عندما تكون هذه التوصيات غير متاحة نظراً لنسبة تشابه ضئيلة أو عدم توفر تلك التوصيات بواسطة النموذج الأول، عندها يقوم بجلب التوصيات بواسطة النموذج الثاني وعرضها للمستخدم.

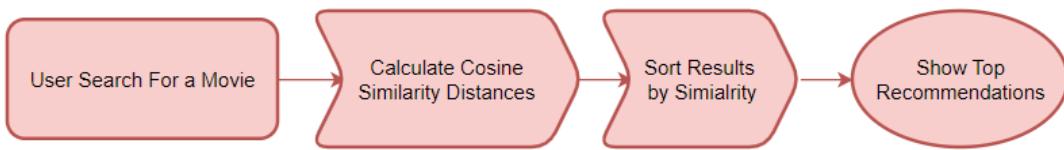


الشكل 10: المخطط التدفقى لنهج التوصية الهجين

### نهج التوصية الأول

تم اتباع نمط بسيط في هذا النموذج، بحيث يقوم باستخدام النموذج المدرب من خلال تطبيق لـ **Cosine Similarity** لتوصية الأفلام المماثلة لفيلم معين.

بحيث يقوم بتحديد الأفلام الأكثر تشابهاً بناءً على إدخال المستخدم أو الفيلم الحالي وترتيبها بحسب درجة التشابه ومن ثم التوصية بالأفلام بناءً على درجة التشابه.



**الشكل 11: مراحل نموذج التوصية الأول**

### نهج التوصية الثاني

في هذا النهج يتم تدريب نموذج توصية هجين يجمع بين التصفية التعاونية والتصفية القائمة على المحتوى.

حيث يقوم بالمراحل التالية:

1. إزالة الكلمات والأحرف التي ليست ذات معنى مثل ( the, a, an, of, in, on, ) (etc.

2. البحث عن أفضل مطابقة ما بين فيلم معين وبيانات الأفلام في النموذج المدرّب مسبقاً بواسطة خوارزمية Cosine Similarity.

3. ترتيب النتائج والتركيز على أكثر 100 نتائجة توصية تم حسابها بواسطة تشابه جيد التمام.

4. حساب متوسط التقييم نسبة إلى عدد الأصوات (المتوسط المرجح Weighted Average) .(Average

$$\text{Weighted Average (WR)} = (v / (v+m)) * R + (m / (v+m)) * C$$

حيث: R هو متوسط التقييم للفيلم

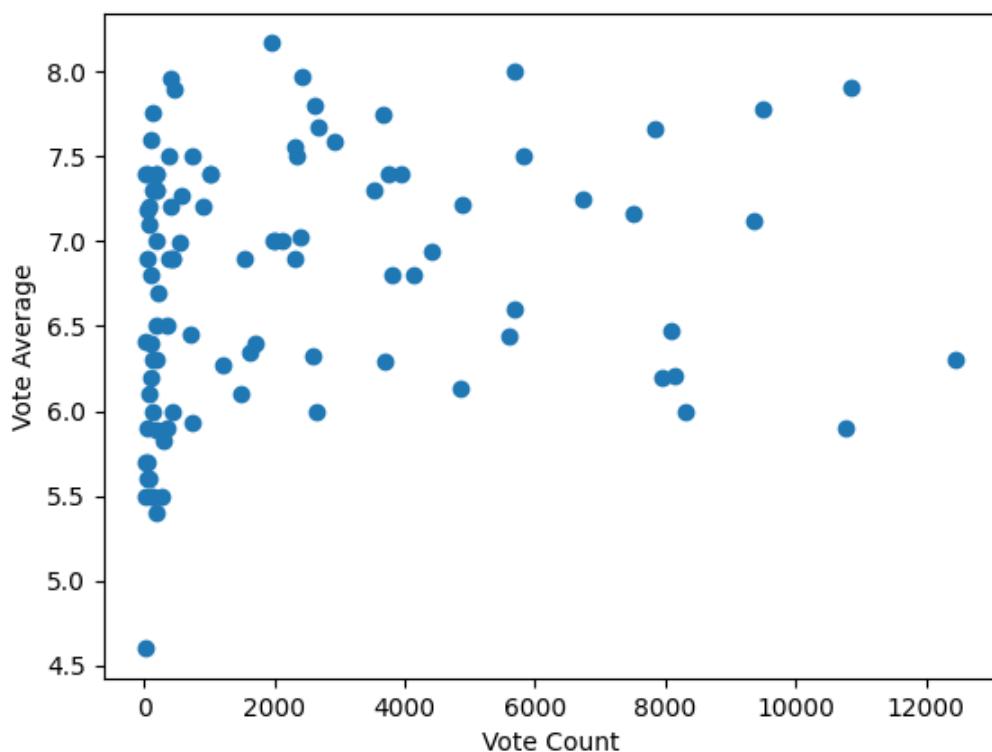
v هو عدد التقييمات التي حصل عليها الفيلم

m هو عدد الأصوات الدنيا التي يمكن أن تؤخذ بعين الاعتبار

C هو متوسط التقييم العام لجميع الأفلام

هذا المعدل يسمح بتقدير الأفلام بناءً على مزيج من متوسط التقييم وعدد الأصوات، حيث يعكس التقييمات الأكثر شيوعاً والمرتبطة بكمية أكبر من الأصوات بشكل أكبر. (24)

علمًا أنه تم حساب اللجوء إلى هذا الحساب لوجود أفلام كثيرة في مجموعة البيانات ذات تقييم مرتفع رغم عدم وجود عدد الأصوات التي قامت بتقدير فيلم معين أو وجود عدد توصيات قليلة كما في الشكل (12). وبالتالي لا يمكن الاعتماد على هذه التقييم المرتفع لوحده بمعزل عن العوامل الأخرى.



الشكل 12: متوسط الأصوات إلى عدد الأصوات

الكود التالي يوضح التعليمات البرمجية المستخدمة في حساب المتوسط المرجح

:Weighted Average

```
weight_averaged_method = second_filters[['id", "title_x", "vote_average", "vote_count"]]

datasets.head(2)

# ### Average rating of the movie
weight_averaged_method['R'] = weight_averaged_method.vote_average
weight_averaged_method.R
# ### Mean vote (or rating) across all movies.
C = weight_averaged_method.R.mean()
C

# ### Number of votes (or ratings) received by the movie.
weight_averaged_method['v'] = weight_averaged_method.vote_count
weight_averaged_method.v
# ### Minimum votes (or ratings) required for the movie to be considered.
PERCENTAGE = .95
m = weight_averaged_method.v.quantile(PERCENTAGE)
print(f"{PERCENTAGE * 100}%", m)
# ### Calculate the weighted average ratings using the provided formula.
weight_averaged_method.head(2)

def WR(query):
    R = query.R
    v = query.v
    return (v / (v + m)) * R + (m / (v + m)) * C

weight_averaged_method["WR"] = weight_averaged_method.apply(WR, axis=1)

weight_averaged_method.head(2)

weight_averaged_method = weight_averaged_method.sort_values("WR", ascending=False)
weight_averaged_method.reset_index(drop=True, inplace=True)
```

5. تصفية الأفلام بناءً على تشابه المحتوى باستخدام Overview ملخص الفيلم لـ

أكثر 50 توصية تشابهاً ناتجة عن المراحل السابقة.

6. يتم استخدام تقنية TF-IDF (التردد النسبي للمصطلح وتردد العكس للمستند)

لتحويل النصوص إلى مصفوفة من الأرقام. تقوم هذه التقنية بتمييز أهمية

الكلمات في النص بناءً على كيفية تكرارها في المستند وتكرارها في مستندات

أخرى. ينتج عن ذلك مصفوفة TF-IDF التي تُستخدم كسمات لتدريب النموذج.

```

tfidf = TfidfVectorizer(stop_words='english')
tfidf_features = tfidf.fit_transform(contentbased.overview_cleaned)

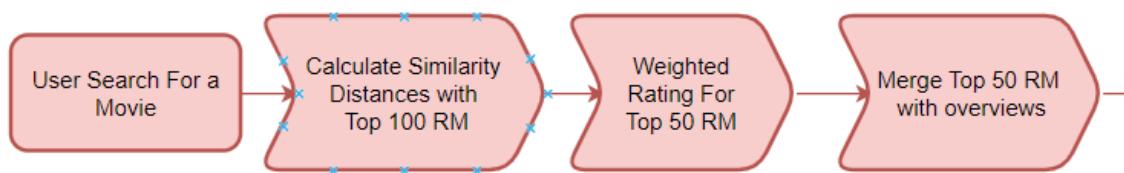
tfidf_features.toarray()

```

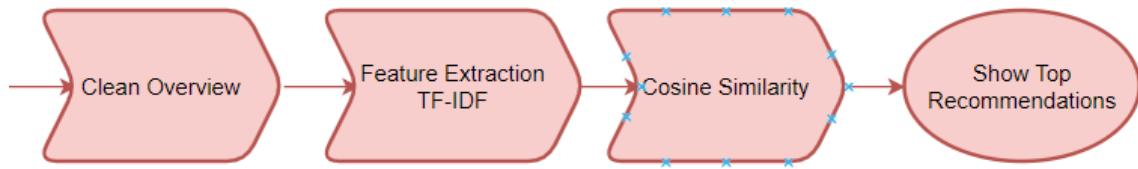
7. ثم يتم تطبيق قياس تشابه جيب التمام على السمات المستخرجة بواسطة-TF-

.IDF

8. وأخيراً، ترتيب النتائج حسب درجة التشابه وإظهار التوصيات للمستخدم.



الشكل 13: مراحل نموذج التوصية الثاني (القسم الأول)



الشكل 14: مراحل نموذج التوصية الثاني (القسم الثاني)

### 4-3-2-3- اختبار النموذج

مثال من خلال الطلب من النموذج اقتراح توصيات على فيلم "The Gentlemen"

```
print(recommend2('The Gentlemen',20))
```

بعد المعالجة الأولية، استخراج أكثر 100 عنصر مشابه لهذا الفيلم من النموذج المدرب مسبقاً بواسطة نسبة تشابه جيد التمام:

وبما أننا قمنا باختيار النسبة المئوية لعدد الأصوات المستخدمة لاقتراح الأفلام بنسبة 95%. وبالتالي، نجد أن الحد الأدنى لعدد الأصوات هو 1039.4 صوتاً.

```
95.0% 1039.4000000000035
```

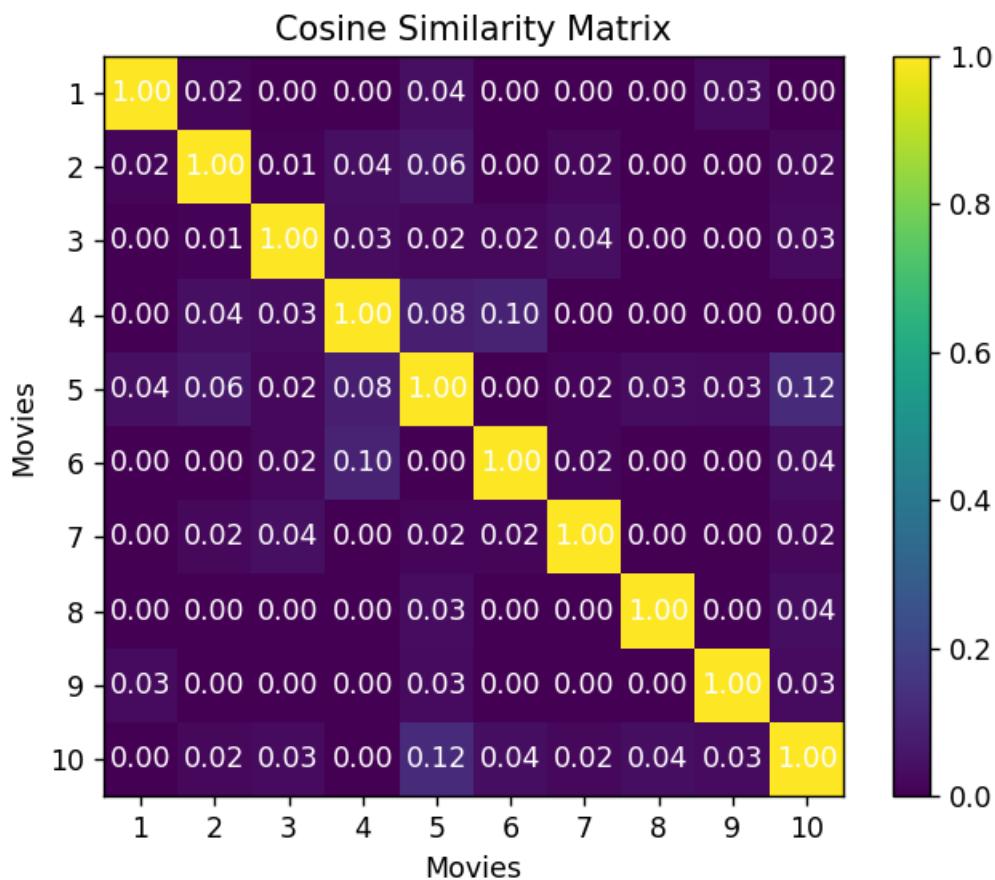
استخراج أكثر 50 عنصر مشابه ودمجها مع ملخص الفيلم Overview، تصفية البيانات الناتجة ومن ثم تطبيق خوارزمية TF-IDF لاستخراج السمات من البيانات الأخيرة.

أخيراً، تطبيق خوارزمية تشابه جيد التمام Cosine Similarity على السمات الناتجة من TF-IDF وترتيب أكثر 20 عنصر تشابهاً لفيلم "The Gentlemen" واختيار أول 8 توصيات لإظهارها للمستخدم.

top_index	cosine_similarity
0	1.000000
1	0.064557
2	0.041023
3	0.030045
4	0.023403
5	0.023016
6	0.022718
7	0.022242
8	0.017285
9	0.011982
10	0.010597
11	0.000000
12	0.000000
13	0.000000
14	0.000000

[ 'Gentlemen Prefer Blondes', 'Dil To Pagal Hai', 'Battle of the Year', 'Hero No. 1', 'Gigi', 'Roberta', 'The Pirate', 'On an Island with You' ]

و مصفوفة التشابه الناتجة بالشكل التالي:



الشكل 15: مصفوفة التشابه لـ *The Gentlemen*

### 3-2-3-5- اختبار دقة التوصيات

لتقييم أداء نظام التوصية الخاص بنا ولقياس مدى كفاءة نظامي التوصية المستخدمين لدينا ومدى تفوق أحدهما على الآخر ، يمكننا اللجوء إلى أحد المقاييس الشائعة في أنظمة التوصية. وفيما يلي بعضًا منها:

مقياس دقة التنبؤات (RMSE): مقياس مستخدم بشكل واسع لقياس دقة التنبؤات النموذج. حيث يحسب الجذر التربيعي لمربع متوسط الفرق بين التقييم المتوقع والتقييم الحقيقي. بحيث تشير قيمة RMSE أقل إلى دقة أعلى. (27)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_i - \check{r}_i)^2}$$

حيث:

- $N$  هو العدد الإجمالي للتنبؤات.
- $r_i$  هو التقييم الحقيقي.
- $\check{r}_i$  هو التقييم المتوقع.

(MAE) Mean Absolute Error: يقيس طولية المتوسط للفروق بين التقييمات المتوقعة والحقيقة. حيث يوفر طريقة أخرى لقياس نسبة الخطأ.

$$MSE = \frac{1}{N} \sum_{i=1}^N |r_i - \check{r}_i|$$

Precision و Recall و F1-Score: هذه المقاييس شائعة في التعامل مع عائدات المعلومات. Precision تقيس دقة التنبؤات الإيجابية، في حين أن Recall تقيس القدرة على الحصول على توقعات ذات صلة، و F1-Score هو مقياس يدمج مفهوم المقاييس السابقين Precision و Recall.

بالإضافة إلى معايير أخرى يمكن من خلالها تقييم النتائج التي نحصل عليها من النموذج المستخدم، مثل مدى تعطية التوصيات لمحتوى الموضع وتنوع التوصيات بالإضافة إلى الحادثة في الحصول على توصيات جديدة أو غير متوقعة ويمكن أن تجذب المستخدم وبالتالي تجنب الملل في عرض توصيات متكررة.

**تطبيق مقياس RMSE في قياس دقة التوصيات لعينة عشوائية:**

للتتأكد بأن التوصيات التي يتم عرضها للمستخدم هي توصيات فعالة وذات صلة بما قد يرغب المستخدم بمشاهدته لاحقاً، قمنا باستخدام مقياس RMSE. وذلك من خلال:

- قياس RMSE لمجموعة عشوائية تم التوصية بها من خلال النموذج الأول (نظام التصفية القائم على المحتوى).
- قياس RMSE لمجموعة عشوائية تم التوصية بها من خلال النموذج الثاني (نظام هجين يجمع بين التصفية التعاونية والتصفية القائمة على المحتوى).
- مقارنة النتائج.

من خلال القيام بالمراحل السابقة تبين أن النموذج الثاني (الهجين) تفوق على النموذج الأول (القائم على المحتوى لوحده) في أغلب الاختبارات:

- النموذج الهجين غالباً ما كان يوفر التوصيات بكفاءة أكثر في حين أن النموذج القائم على المحتوى لم يقدم نتائج في بعض الاختبارات خاصة في الحالات التي لا تتوفر فيها معلومات كافية عن المحتوى.
- النموذج الهجين يمر بعدة مراحل تتضمن تنظيف للبيانات وعدة طرق لحساب التشابه وإعادة ترتيب النتائج وانتخاب أكثر العناصر تشابهاً لذلك كانت دقة التوصيات الناتجة أعلى بنسبة تتراوح ما بين 5% إلى 35% أكثر من النموذج القائم على المحتوى.

أمثلة:

Title	Model 1	Model2
Meg 2: The Trench	0.404508	0.261136
Scream VI	0.459374	0.335098
Spider-Man: No Way Home	0.357390813	0.146985794
Encanto	0.340693	0.318162
The Notebook	0.428682	0.291432
Twilight	0.502888	0.422228

الجدول 2: أمثلة تفوق النموذج الهجين على النموذج القائم على المحتوى

- والجدير بالذكر أيضاً أن النموذج القائم على المحتوى قدم نتائج أكثر دقة في بعض الاختبارات، ويعود ذلك إلى توفر معلومات كافية عن المحتوى وغياب

تقييمات المستخدمين (عدد التقييمات/ التصويتات، متوسط التقييم/ التصويت أو كلاهما).

أمثلة:

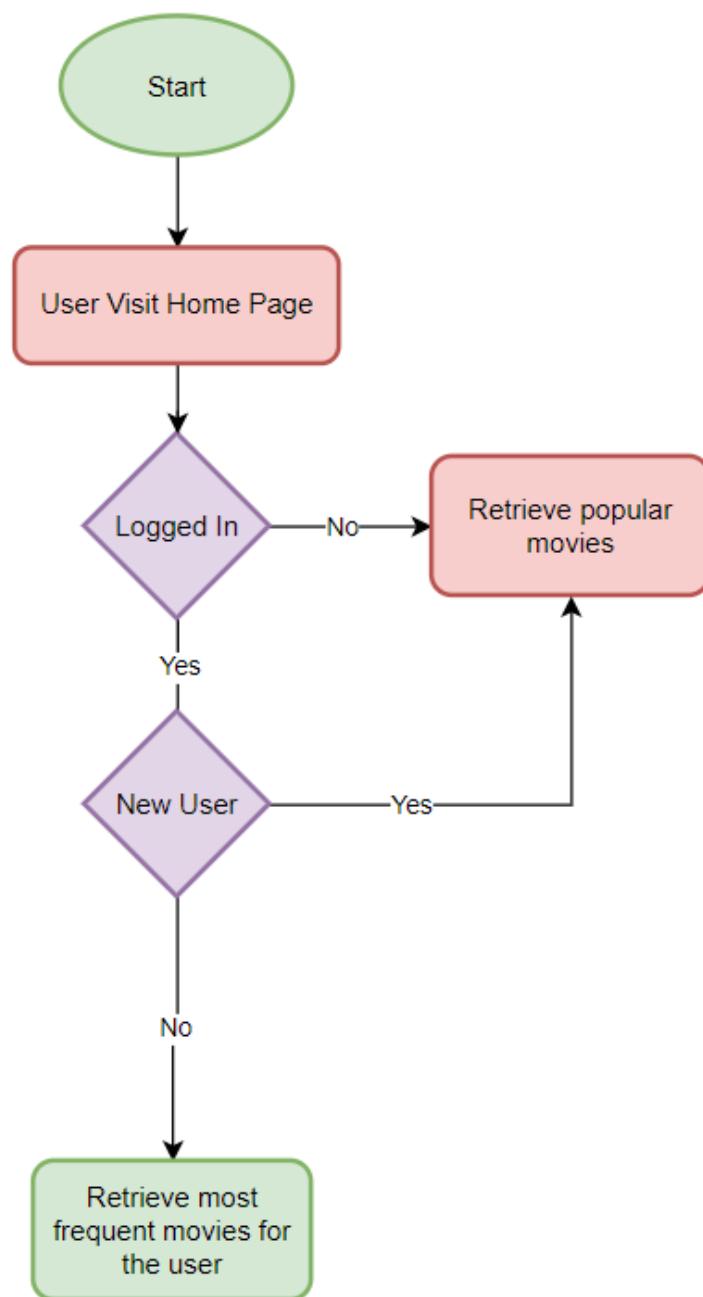
Title	Model 1	Model2
Barbie	0.391182	0.572321
Frozen	0.510298	0.693371
The Valet	0.427186	0.446765
Iron Man	0.332073	0.345484

الجدول 3: أمثلة تفوق النموذج القائم على المحتوى على النموذج الهجين

وبالتالي ومن خلال الدمج بين النموذجين توصلنا لتوصيات أكثر فعالية، فعندما يكون النظام القائم على المحتوى قادرًا على تقديم توصيات بدقة عالية ومتاحة في الزمن الحقيقي. عندها يتم إظهار التوصيات للمستخدم بناءً على نتائج توصيات النموذج الأول. وعندما يتذرع الحصول على توصيات دقيقة في الزمن الحقيقي وعندما يعتذر على النموذج إيجاد نتائج مشابهة يتم اللجوء إلى النموذج الهجين للحصول على التوصيات المناسبة وعرضها للمستخدم.

وهكذا أصبح نظام التوصية الخاص بنا جاهزًا لتنقل إلى موقع الأفلام الذي سنهم ببنائه بإظهار تلك التوصيات للمستخدم بواجهات جذابة ومفهومة.

بناءً على مخطط حالات الاستخدام السابقة والدراسة التحليلية السابقة. بالإضافة إلى شرح نظام التوصية الذي قمنا ببنائه لخضير تجربة المستخدم. يمكن تلخيص ذلك في المخطط التدفق التالي:



الشكل 16: المخطط التدفق لإظهار البيانات للمستخدمين

## 4- الفصل الرابع: النتائج

### 4-1 واجهات الويب

The screenshot shows the LMovies website with a dark purple header. The header includes the LMovies logo, navigation links for Discover, About, and Contact Us, a search bar labeled "Search LMovies DB", and a "Login" button. The main content area has a light purple background. It features a section titled "About Us" with a brief welcome message, a mission statement, and a thank you note. Below this is a "Project Creator Details" section listing the creator's name, username, university, and degree. At the bottom of the page is a footer with links for Discover, About, Legal, and Follow Us, along with a logo for Syrian Virtual University.

**About Us**

Welcome to Movie Recommender, your go-to platform for discovering new and exciting movies tailored just for you. Our mission is to make your movie-watching experience enjoyable and hassle-free. We use advanced algorithms to recommend movies based on your preferences, ensuring that you never miss a great film. Thank you for choosing Movie Recommender. Sit back, relax, and let us guide you through the world of cinema!

**Project Creator Details**

**Name:** Lina Mhd Imad Mouselli  
**Username:** Lina\_114886  
**University:** Syrian Virtual University  
**Degree:** Master in Web Technologies

**DISCOVER**  
Search

**ABOUT**  
About LMovies

**LEGAL**  
Terms of Service

**FOLLOW US**  
[f](#)

## 4-1-1- واجهات الويب للزائر

### الصفحة الرئيسية

The screenshot shows the LMovies homepage with a dark purple header. The header includes the LMovies logo, navigation links for Discover, About, Contact Us, a search bar labeled "Search LMovies DB", and a "Login" button. Below the header, a "Welcome" message is displayed, followed by a paragraph about the project's affiliation with Syrian Virtual University. A section titled "Trending Movies" features posters for several movies: Inception, Pulp Fiction, Avengers: Infinity War, The Dark Knight, Iron Man, Deadpool, Guardians of the Galaxy, and Interstellar.

### صفحة البحث

The screenshot shows the LMovies search page with a dark purple header. The header includes the LMovies logo, navigation links for Discover, About, Export Dataset, a search bar labeled "Search LMovies DB", and a "Sign Up" button. The main content area is titled "Search Movies". It features search filters for Title (set to "Twilight"), Start date (set to "1990-01-01"), End date (set to "2023-01-01"), and Genres (a dropdown menu showing All Genres, Action, Adventure, Animation). A "Search" button is located below the filters. At the bottom, there are pagination controls for "Page 1 of 21", "Next", and "»". A table displays movie results for "The Twilight Saga: Breaking Dawn - Part 2", including columns for #, Title, Genres, Original Language, Production Companies, Release Date, Vote Count, Poster, and Action. The "Action" column for the first result contains a "Show" button.

## صفحة تفاصيل الفيلم

LMovies  
A World of Movie Wonders

Discover   About   Export Dataset

Search LMovies DB

Sign Up

### The Twilight Saga: Eclipse

genres	Adventure-Fantasy-Drama-Romance
original_language	en
overview	Bella once again finds herself surrounded by danger as Seattle is ravaged by a string of mysterious killings and a malicious vampire continues her quest for revenge. In the midst of it all she is forced to choose between her love for Edward and her friendship with Jacob knowing that her decision has the potential to ignite the ageless struggle between vampire and werewolf. With her graduation quickly approaching Bella is confronted with the most important decision of her life.
popularity	77.139
production_companies	Summit Entertainment-Maverick Films-Imprint Entertainment-Sunswept Entertainment-Temple Hill Entertainment
release_date	June 23, 2010
budget	68000000.0
revenue	698491347.0
runtime	124
status	Released
tagline	It all begins... with a choice.
vote_average	6.2
vote_count	7938
credits	Kristen Stewart-Robert Pattinson-Taylor Lautner-Bryce Dallas Howard-Dakota Fanning-Billy Burke-Peter Facinelli-Elizabeth Reaser-Jackson Rathbone-Kellan Lutz

#### Cast Images and Details

Kristen Stewart  
4 - Bella Swan

Robert Pattinson  
5 - Edward Cullen

Taylor Lautner  
6 - Jacob Black

Billy Burke  
7 - Charlie Swan

Peter Facinelli

Elizabeth Reaser

Jackson Rathbone

### Recommended Movies



**The Twilight Saga: New Moon**  
Release Date: Nov. 18, 2009  
Average: 6.0



**The Twilight Saga: Breaking Dawn - Part 1**  
Release Date: Nov. 16, 2011  
Average: 6.206



**Twilight**  
Release Date: Nov. 20, 2008  
Average: 6.3



**The Twilight Saga: Breaking Dawn - Part 2**  
Release Date: Nov. 13, 2012  
Average: 6.474

مثال آخر



#### Interstellar

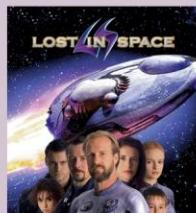
genres	Adventure-Drama-Science Fiction
original_language	en
overview	The adventures of a group of explorers who make use of a newly discovered wormhole to surpass the limitations on human space travel and conquer the vast distances involved in an interstellar voyage.
popularity	151.161
production_companies	Legendary Pictures-Syncoopy-Lynda Obst Productions
release_date	Nov. 5, 2014
budget	165000000.0
revenue	701729206.0
runtime	169
status	Released
tagline	Mankind was born on Earth. It was never meant to die here.
vote_average	8.391
vote_count	30708
credits	Matthew McConaughey-Anne Hathaway-Jessica Chastain-Mackenzie Foy-Michael Caine-Bill Irwin-Casey Affleck-John Lithgow-Wes Bentley-David Gyasi-Matt Damon-Topher Grace-Timothée Chalamet-Ellen Burstyn-David Oyelowo-William Devane-Josh Stewart-Collette Wolfe-Leah Cairns-Russ Fega-Lena

### Cast Images and Details



**Matthew McConaughey**  
9 - Joseph "Coop" Cooper

### Recommended Movies



**Lost in Space**

Release Date: April 3, 1998

Average: 5.4

**The Midnight Sky**

Release Date: Dec. 10, 2020

Average: 5.8

**Space Chimps**

Release Date: June 25, 2008

Average: 4.8

**Stargate: Continuum**

Release Date: July 29, 2008

Average: 7.1



**Planet of the Apes**

Release Date: Feb. 7, 1968

Average: 7.642



**Planet of the Apes**

Release Date: July 25, 2001

Average: 5.725



**Salyut-7**

Release Date: Sept. 22, 2017

Average: 7.1

## 4-1-2- واجهات الويب للتسجيل وتسجيل الدخول

### صفحة التسجيل

The screenshot shows the LMovies website's registration page. At the top, there is a dark header bar with the LMovies logo, a search bar labeled "Search LMovies DB", and a "Login" button. Below the header, the URL "Home / Register" is visible. The main form has three input fields: "Username", "Password", and "Confirm Password", all of which are redacted with black bars. Below the form is a small text block stating: "By clicking register you are confirming you have read our Terms of Service, agree to all our policies, and understand these may both change at any time without advance warning." A "Register" button is located at the bottom of the form.

DISCOVER

ABOUT

LEGAL

FOLLOW US

LMovies © 2023 LMovies.com ®. ALL RIGHTS RESERVED.

### صفحة تسجيل الدخول

The screenshot shows the LMovies website's login page. At the top, there is a dark header bar with the LMovies logo, a search bar labeled "Search LMovies DB", and a "Login" button. Below the header, the URL "Home / Login" is visible. To the right of the URL, there is a "Create an Account" button. The main form has two input fields: "username" and "Password", both redacted with black bars. Below the form is a "Login" button. At the bottom of the page, there is a footer bar with links to "Discover", "About", "Contact Us", "Terms of Service", and "Privacy Policy".

DISCOVER

ABOUT

LEGAL

FOLLOW US

LMovies © 2023 LMovies.com ®. ALL RIGHTS RESERVED.

## 4-1-3- واجهات الويب للمستخدم

### الصفحة الرئيسية

The screenshot shows the LMovies website interface. At the top, there's a dark header bar with the LMovies logo, a search bar containing "Search LMovies DB", and a "Logout, Emad" button. Below the header, a welcome message "Welcome Emad" is displayed, followed by a paragraph of text about the project's affiliation with Syrian Virtual University. A section titled "Recommended for You" features a grid of movie posters for "THE NUN", "BLACK HAWK DOWN", "THE MEG", "STAR WARS: THE LAST JEDI", and four other movies whose titles are not clearly legible. Below this, there are two rows of movie posters: "THE NUN", "THE EXORCIST", "RISE", and "THE CONJURING". Further down, a section titled "Trending Movies" displays posters for "INTERSTELLAR", "INCEPTION", "IRON MAN", and "AVATAR". At the bottom of the page, there are additional movie posters for "Avengers: Endgame", "PULP FICTION", "Forrest Gump", and "The Dark Knight".

## صفحة سجلات البحث الأخيرة

The screenshot shows a table titled 'My Last Searches' with columns: Title, Start Date, End Date, Genres, and Search Date. The data includes various movie titles like 'The Gentlemen', 'Tangled', and 'Spider Man', along with their search dates and genres where applicable.

Title	Start Date	End Date	Genres	Search Date
The Gentlemen	-	-	-	23-January-2024 05:18 PM
Tangled	-	-	-	20-January-2024 01:06 PM
-	-	-	Horror	20-January-2024 12:28 PM
Home Alone	-	-	-	20-January-2024 10:18 AM
Hangover	-	-	-	20-January-2024 10:12 AM
Hungover	-	-	-	20-January-2024 10:11 AM
Dump and Dumper	-	-	-	20-January-2024 10:10 AM
lier lier	-	-	-	20-January-2024 10:09 AM
Scream	-	-	-	20-January-2024 10:08 AM
Home Alone	-	-	-	20-January-2024 10:04 AM
Twilight	-	-	-	20-January-2024 10:03 AM
Scream	-	-	-	20-January-2024 09:39 AM
Batman	-	-	-	19-January-2024 08:18 PM
Spider Man	-	-	-	19-January-2024 08:17 PM

## 4-1-4- واجهات المدير

### صفحة المستخدمين

The screenshot shows a table titled 'Users' with columns: Username and Action. It lists three users: Emad, Limar, and Mohammed, each with a 'Delete' link next to their name.

Username	Action
Emad	Delete
Limar	Delete
Mohammed	Delete

#### DISCOVER

Search  
Search History

#### ABOUT

About LMovies  
Contact Us

#### LEGAL

Terms of Service  
Privacy Policy

#### FOLLOW US

LMovies © 2023 LMovies.com ®. ALL RIGHTS RESERVED.

95

## صفحة المدير

### مرتبة حسب سنة الإصدار من الأحدث إلى الأقدم

The screenshot shows the LMovies Admin Panel's Statistics section. At the top, there are links for Discover, About, Admin Panel, and a search bar labeled "Search LMovies DB". On the right, it says "Logout, admin". Below the header, it displays "Total Movies: 662084" and "Total Users: 4". A heading "Statistics" is followed by a table titled "Movies Per Year". The table lists years from 2049 down to 2015, showing the count of movies for each year. At the bottom of the table, there are navigation links: "Page 1 of 10", "Next", and "»".

Year	Movie Count
2049	1
2028	1
2027	1
2026	3
2025	15
2024	75
2023	1532
2022	10923
2021	30057
2020	30028
2019	30443
2018	27671
2017	26208
2016	23236
2015	22324

### مرتبة حسب عدد الأفلام من الأكثر إلى الأصغر

The screenshot shows the LMovies Admin Panel's Statistics section. At the top, there are links for Discover, About, Admin Panel, and a search bar labeled "Search LMovies DB". On the right, it says "Logout, admin". Below the header, it displays "Total Movies: 662084" and "Total Users: 4". A heading "Statistics" is followed by a table titled "Movies Per Year". The table lists years from 2019 down to 2007, showing the count of movies for each year. At the bottom of the table, there are navigation links: "Page 1 of 10", "Next", and "»".

Year	Movie Count
2019	30443
2021	30057
2020	30028
2018	27671
2017	26208
2016	23236
2015	22324
2014	21607
2013	19667
2012	16490
2011	13809
2010	12697
2009	12031
2008	11300
2007	10981

## صفحة إدخال البيانات

The screenshot shows the LMovies Admin Panel with a dark purple header. The header includes the LMovies logo, navigation links for Discover, About, and Admin Panel, a search bar for 'Search LMovies DB', and a 'Logout, admin' link. The main content area has a light purple background and displays the 'Import Movies' section. It asks for a CSV file upload and provides a file input field with a 'Browse ...' button and an 'Upload' button. At the bottom of the page is a dark footer with links for Discover, About, Legal, and Follow Us, along with the LMovies logo and copyright information.

## طلب تصدير البيانات

This screenshot shows the same Admin Panel interface as the previous one, but the 'Import Movies' section is now expanded to include 'Export Dataset' and 'Import Dataset' buttons. The rest of the page, including the footer, remains identical to the first screenshot.

طلب كلمة المرور من المدير للوصول إلى البيانات:

The screenshot shows the LMovies Admin Panel interface. At the top, there's a navigation bar with 'Discover', 'About', 'Admin Panel' (selected), and a search bar 'Search LMovies DB'. On the right, it says 'Logout, admin'. Below the navigation, there's a section titled 'Import Movies' with a sub-section 'Please Upload Movies as CSV File:'. It features a file input field with a placeholder 'Browse ...', a 'Upload' button, and a 'CSV File' link. To the right of this is a sidebar with 'Dashboard', 'Users', 'Export Dataset' (selected), and 'Import Dataset'. A modal window is open in the center, prompting for a password with the text 'Please enter the password:' and a 'Send' button. The footer contains links for 'DISCOVER' (Search, Search History), 'ABOUT' (About LMovies, Contact Us), 'LEGAL' (Terms of Service, Privacy Policy), and 'FOLLOW US' (Facebook icon). The copyright notice '© 2023 LMovies.com ®. ALL RIGHTS RESERVED.' is also present.

ثم يتم تنزيل البيانات ك ملف لاحقته "csv"



## 4- الفصل الخامس: الآفاق المستقبلية

- زيادة حجم وتوع ببيانات المستخدم لتعزيز دقة التوصيات وتعزيز فهم احتياجات المستخدمين وتقديراتهم.
- تضمين خوارزميات التوصية القائمة على السياق للحصول على توصيات أوسع وأكثر ملائمة للسياق الخاص بالمستخدم.
- تمكين المستخدمين من إعطاء تقييمات وأراء حول الأفلام.
- إضافة عروض الأفلام لكي يتمكن المستخدم من مشاهدة الأفلام عبر الموقع.

## الخلاصة Conclusion

تُعد هذه الدراسة تأكيداً على أهمية مجال نظم التوصية، وأهمية العمل على تطوير تقنيات متقدمة في هذا المجال، خاصةً على منصات عرض الأفلام، حيث يشكل تحدي تقديم تجربة مستخدم فعالة ومثيرة في ظل تزايد العروض والمحظى المتاح.

قدمنا في هذه الدراسة مقترن لتحليل سلوك المستخدم أثناء استخدامه لموقع الأفلام ومن خلال سجل بحثه لتقديم مقترنات متخصصة تلبي تطلعاتهم باستخدام نظام توصية هجين يجمع بين نظام التصفية التعاونية ونظام التصفية القائم على المحتوى. حيث تم استخدام مجموعة بيانات الأفلام الخاصة بـ TMDB بعد سجلات حوالي 700 ألف فيلم بمختلف التصنيفات والتقييمات وبفترة زمنية من الثمانينات حتى التاريخ الحالي. كما تم استخدام ADI الخاصة بـ TMDB لعرض صور البوستر الخاص بالفيلم وصور الممثلين وتفاصيلهم. قمنا بذلك باستخدام ثلاثة خوارزميات في نظام التوصية الهجين (Cosine Similarity, TF-IDF, Weighted Average). كما تم بناء واجهات ويب لكل من الزائر والمستخدم ومدير الموقع بسيطة وسهلة التعامل.

في الختام، يُظهر البحث أهمية تحسين نظم التوصية في مجال عرض الأفلام وكيف يمكن أن تلعب هذه التقنيات دوراً حيوياً في تحسين تفاعل المستخدمين وزيادة جاذبية المحتوى. تم تحقيق أهداف البحث من خلال تصميم نظام توصية متقدم يتناسب مع احتياجات المستخدمين ويساهم في تحسين تجربتهم.

## المراجع

1. Dong, Z., Wang, Z., Xu, J., Tang, R., & Wen, J. (2022). *A Brief History of Recommender Systems*.
2. Bhave, A., Kulkarni, H., Biramane, V., & Kosamkar, P. (2015, January). *Role of different factors in predicting movie success*. In *2015 International conference on pervasive computing (ICPC)* (pp. 1-4). IEEE
3. Bobadilla, J., Ortega, F., Hernando, A., & García-Serrano, A. (2013). *A Survey of Recommender Systems*. *Knowledge-Based Systems*, 46, 109-132.
4. Burke, R. (2002). *Hybrid recommender systems: Survey and experiments*. *User modeling and user-adapted interaction*, 12(4), 331-370.
5. Soanpet, Sree Lakshmi, & Dr. T. Adi Lakshmi. (2014). *Recommendation Systems: Issues and Challenges*. Dept. of CA, Vasavi College of Engineering, Hyderabad, India.
6. He, X., Liao, L., Zhang, H., Li, L., & Chua, T. S. (2017). *Deep Learning based Recommender System: A Survey and New Perspectives*.
7. Su, X., & Khoshgoftaar, T. M. (2009). *A Survey of Collaborative Filtering Techniques*.
8. Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). *Recommender Systems: An Introduction*.

9. Koren, Y., Bell, R., & Volinsky, C. (2009). *Matrix Factorization Techniques for Recommender Systems*.
10. Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*.
11. Ricci, F., Rokach, L., & Shapira, B. (2015). *Introduction to recommender systems handbook*. Springer.
12. Adomavicius, G., & Tuzhilin, A. (2005). *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*, 17(6), 734-749.
13. Konstan, J. A., & Riedl, J. (2012). *Recommender systems: from algorithms to user experience*. *User modeling and user-adapted interaction*, 22(1-2), 101-123.
14. Oard, D. W., & Resnik, P. (Eds.). (2013). *Information retrieval and the web* (Vol. 53). Springer Science & Business Media.
15. Zhang, J., Chowdhury, A., & Liu, Y. (2019). *A review of deep learning-based recommendation approaches and techniques*. *Journal of Big Data*, 6(1), 1-39.
16. Wang, Y., Wang, X., Liu, X., & Wang, C. (2018). *Collaborative filtering recommendation algorithm based on matrix factorization*. *Journal of Physics: Conference Series*, 1006(1), 012008.
17. Singh, M., & Gupta, A. (2018). *Recommendation systems: A review*. *International Journal of Computer Applications*, 180(34), 1-8.

18. Su, X., & Khoshgoftaar, T. M. (2009). *A survey of collaborative filtering techniques*. *Advances in Artificial Intelligence*, 2009, 421425.
19. Wang, C., & Blei, D. M. (2011). *Collaborative topic modeling for recommending scientific articles*. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 448-456.
20. "Collaborative Filtering for Implicit Feedback Datasets" by Y. Hu, Y. Koren, and C. Volinsky
21. "Content-Based Recommender Systems: State of the Art and Trends" by L. Chen, X. Wang, and T. Qin
22. Netflix Prize. 2009. *The Netflix Prize*. Retrieved December 6, 2015 from <http://www.netflixprize.com/>.
23. Netflix Technology Blog. "Netflix Recommendations: Beyond the 5 Stars." *Netflix Technology Blog*, URL: <https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars>.
24. Omega, C. Z. (2021, September). *Movie Recommendation System Using Weighted Average Approach*. In *2021 2nd International Conference on Innovative and Creative Information Technology (ICITech)* (pp. 105-109). IEEE.
25. Zhou, R., Khemmarat, S., & Gao, L. (2010, November). *The impact of YouTube recommendation system on video views*. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement* (pp. 404-410).

26. Musa, J. M., & Zhihong, X. (2020, April). *Item based collaborative filtering approach in movie recommendation system using different similarity measures*. In *Proceedings of the 2020 6th International Conference on Computer and Technology Applications* (pp. 31-34).
27. Avazpour, I., Pitakrat, T., Grunske, L., & Grundy, J. (2014). *Dimensions and metrics for evaluating recommendation systems*. *Recommendation systems in software engineering*, 245-273.
28. "Coursera." Coursera, <https://www.coursera.org>.
29. Piotte, M., & Chabbert, M. (2009). *The pragmatic theory solution to the netflix grand prize*. Netflix prize documentation.
30. "Carleton College Recommender Systems." Carleton College Computer Science Comps,  
[https://www.cs.carleton.edu/cs\\_comps/0607/recommend/recommender/index.html](https://www.cs.carleton.edu/cs_comps/0607/recommend/recommender/index.html).
31. Real Python. "Building a Recommendation Engine with Collaborative Filtering." Real Python, <https://realpython.com/build-recommendation-engine-collaborative-filtering/>
32. OpenML. OpenML, <https://www.openml.org>.
33. Kaggle. Kaggle, <https://www.kaggle.com>.
34. Stack Overflow Community. Stack Overflow,  
<https://stackoverflow.com/>.
35. YouTube. YouTube, [https://www.youtube.com/](https://www.youtube.com)
36. Netflix. Netflix, [https://www.netflix.com/](https://www.netflix.com)