

تأثير السعر على جودة أنظمة التوصية في التجارة الإلكترونية

Apriori و K-Means دراسة تجريبية باستخدام خوارزميات

الملخص

تُعد أنظمة التوصية من الركائز الأساسية في منصات التجارة الإلكترونية، حيث تُستخدم لتحسين تجربة المستخدم، تقليل وقت اتخاذ القرار، وزيادة معدلات التحويل والمبيعات. وبينما تعتمد هذه الأنظمة على تحليل سلوك المستخدمين والخصائص المشتركة بين المنتجات، يُعد متغير "السعر" من بين أكثر السمات إثارة للجدل من حيث تأثيرها على جودة التوصيات.

يهدف هذا البحث إلى تقييم تأثير السعر كمتغير تحليلي في بناء نماذج التوصية، مع التركيز على استخدام خوارزمية لاستخلاص قواعد الارتباط. أُجريت دراسة تجريبية على مجموعة بيانات تضم 25 منتجًا **Apriori** للعنقدة و **K-Means** متنوعًا من حيث الخصائص، إلى جانب بيانات فواتير الشراء.

أظهرت النتائج أن تضمين السعر في نموذج التجميع قد يُحسن التماسك السعري داخل العناقيد، لكنه في المقابل يُضعف التشابه الوظيفي والخصائص الفيزيائية بين المنتجات المُجمعة. بالتالي، فإن تضمين السعر كمُدخل مباشر في خوارزميات التوصية لا يضمن تحسين الجودة، بل قد يُقلل من دقة التوصيات في سياقات معينة، مما يستدعي تقييمًا دقيقًا وتجريبًا متعدد النماذج قبل اعتماده.

العنقدة، قواعد الارتباط، نظم التوصية، التجارة، **Apriori**، **K-Means**، الكلمات المفتاحية: السعر، جودة التوصية الإلكترونية، التنقيب في البيانات.

1. المقدمة

مع تضخم حجم البيانات الرقمية وتنوع المنتجات والخدمات في البيئة الرقمية، أصبحت أنظمة التوصية أدوات حيوية لتصفية المحتوى وتقديم اقتراحات مخصصة للمستخدمين. وتشغل هذه الأنظمة على تحليل أنماط السلوك الشرائي، التشابه بين المنتجات، والتفضيلات السابقة، بهدف توقع ما قد يرغب المستخدم في شرائه مستقبلاً.

رغم تنوع السمات المستخدمة في بناء هذه الأنظمة (مثل النوع، الوزن، الشركة، ونوع الاستخدام)، يُعد **السعر** من أكثر المتغيرات تأثيرًا – بل وأكثرها جدلاً – نظرًا لارتباطه الوثيق بقرار الشراء. ومع ذلك، فإن تضمين السعر كمُدخل كمّي مباشر في النماذج التحليلية قد يؤدي إلى نتائج متضاربة: فمن ناحية، قد يعزز التجميع الدقيق حسب الفئة السعرية، لكنه من ناحية أخرى قد يُضعف التماسك الوظيفي بين المنتجات.

يهدف هذا البحث إلى تحليل تأثير السعر على جودة التوصيات من خلال دراسة تجريبية تستخدم تقنيات **العنقدة** عند تضمين **K-Means** وتحديداً، يتم تقييم أداء خوارزمية **(Association Rules)** وقواعد الارتباط **(Clustering)** لاستخلاص قواعد توصية مبنية على سلوك الشراء. **Apriori** السعر مقابل استبعاده، إلى جانب استخدام خوارزمية وتُقاس جودة التوصيات من خلال تقييم بشري يركز على درجة التشابه الوظيفي بين المنتجات المقترحة.

2. الأسس النظرية

2.1 السعر كمتغير تحليلي

يُعد السعر متغيرًا كمّيًا يعكس القيمة السوقية للمنتج، ويُستخدم غالبًا لتصنيف المنتجات إلى فئات (منخفض، متوسط، مرتفع). في سياق التوصية، يمكن أن يُستخدم السعر لتقسيم المستخدمين أو المنتجات إلى شرائح سلوكية، لكن هيمنته العددية قد تُهمش سمات أخرى أقل تباينًا، مما يُحدث تحيزًا في نتائج التجميع.

2.2 خوارزمية K-Means

عناقيد بناءً k من أشهر خوارزميات التجميع غير الخاضعة للإشراف، وتهدف إلى تقسيم البيانات إلى K-Means تُعد فإن هذه الخوارزمية "تُجري تجميعًا **scikit-learn** على التشابه في المسافات الإقليدية بين النقاط. كما توضح وثائق تلقائيًا للعناصر المتشابهة في مجموعات"، وتُستخدم في تطبيقات مثل تقسيم العملاء وتجميع نتائج التجارب.

يتأثر بشكل كبير بقياسات السمات، مما يجعل تطبيع البيانات ضروريًا، ويُعطي K-Means ومع ذلك، فإن أداء السمات ذات القيم العالية (مثل السعر) وزنًا أكبر إن لم تُعالج.

2.3 خوارزمية Apriori

لاستخلاص قواعد الارتباط من مجموعات المعاملات (مثل فواتير الشراء)، وتُحدد الأزواج Apriori تُستخدم خوارزمية **الثقة (Support)** أو المجموعات التي تُشتري معًا بشكل متكرر. تعتمد هذه القواعد على مقاييس مثل **الدعم** وتُستخدم لاقتراح منتجات مكملة. يمكن دمج هذه القواعد مع نتائج التجميع لتحسين **(Lift)** و**الرفع (Confidence)** دقة التوصية.

2.4 جودة التوصية

(Relevance) والصلة، **(Coverage)** التغطية، **(Precision)** تُقاس جودة نظام التوصية بعدة مقاييس، منها الدقة في هذه الدراسة، يُعتمد على **تقييم بشري** يركز على "درجة التشابه الوظيفي" بين المنتجات المقترحة، حيث يتم تقييم مدى منطقية التوصية من منظور المستخدم النهائي.

3. المنهجية

3.1 مصدر البيانات

تم استخدام مجموعة بيانات تضم 25 منتجًا من فئات متنوعة، تتضمن الخصائص التالية:

- اسم المنتج
- السعر
- الوزن
- النوع (مثلًا: إلكتروني، منزلي، رياضي)
- الشركة المصنعة
- نوع الاستخدام

بالإضافة إلى ذلك، تم جمع بيانات فواتير شراء من مستخدمين وهميين لتحليل أنماط الشراء.

3.2 معالجة البيانات

تم اتباع الخطوات التالية لضمان جودة البيانات:

- **معالجة القيم المفقودة:** استُبدلت القيم الفارغة بمقاييس مركزية (المتوسط أو الوسيط).
- لتحديد وإزالة القيم المتطرفة، خاصة في متغير (IQR) **كشف القيم الشاذة:** استُخدم نطاق التباين الربيعي السعر.
- على جميع السمات الكمية لضمان معاملة متساوية في (Min-Max Scaling) **تطبيع البيانات:** طُبّق التطبيع K-Means خوارزمية.

3.3 التصميم التجريبي

أُجريت التجربة على مرحلتين:

K-Means المرحلة 1: التجميع باستخدام

- **بإدراج السعر** كمتغير تحليلي K-Means التجربة أ: تنفيذ.
- **بدون السعر**، والاعتماد فقط على السمات الوظيفية (الوزن، النوع، إلخ) K-Means التجربة ب: تنفيذ.
- $k = 5$ وتم تثبيته عند (Elbow Method) "تم تحديده باستخدام طريقة "المرفق (k) عدد العناقيد.

المرحلة 2: استخلاص قواعد الارتباط

- لاستخراج قواعد الترابط من بيانات الفواتير (mlxtend) من خلال مكتبة **Apriori** استُخدمت خوارزمية.
- تم تحديد حدود دنيا للدعم (0.1) والثقة (0.5) لتصفية القواعد ذات الدلالة.

المرحلة 3: بناء نظام التوصية

Apriori: تم دمج نتائج التجميع مع قواعد

- إذا كان المستخدم يتصفح منتجًا في عنقود معين، تُقترح عليه منتجات من نفس العنقود.
- ("B" اشترى أيضًا A، مثلًا: "من اشترى") إن وُجدت Apriori تُضاف توصيات من قواعد.

المرحلة 4: التقييم

- تم عرض التوصيات على مجموعة من المقيمين (10 أشخاص).
- كُلفوا بتقييم "درجة التشابه الوظيفي" بين المنتج الأصلي والمنتج المقترح على مقياس من 1 إلى 100.
- حُسب المتوسط العام لكل تجربة.

4. النتائج

أُجريت مقارنة شاملة بين التجربتين، وأظهرت النتائج ما يلي:

المعيار	مع تضمين السعر	بدون تضمين السعر
متوسط تقييم التشابه	72.4 / 100	85.6 / 100
عدد العناقيد	5	5
الوضوح الوظيفي	متوسط	مرتفع
تباين السعر داخل العنقود	منخفض	عالي

التحليل الكمي

- عند تضمين السعر، كانت العناقيد متماسكة من حيث السعر (تباين منخفض)، لكنها ضمت منتجات مختلفة وظيفيًا (مثلًا: "سماعة رأس" مع "بطارية").
- عند استبعاد السعر، تحسّن التشابه الوظيفي بشكل ملحوظ، حيث تجمعت المنتجات حسب النوع والاستخدام (مثلًا: جميع الأجهزة الرياضية في عنقود واحد).
- أضافت قيمة تكميلية، خاصة في حالات الشراء المترافق (مثل: "حقيبة + شاحن") Apriori قواعد

5. المناقشة

تشير النتائج إلى أن السعر، رغم أهميته في اتخاذ القرار الشرائي، قد يحدث تحيزًا في عملية التجميع. فبسبب مما يجعله العامل K-Means، طبيعته الكمية العالية، يُهيمن السعر على المسافات الإقليدية المستخدمة في المحدد الرئيسي في تشكيل العناقيد، حتى لو كانت السمات الوظيفية غير متجانسة.

هذا التحيز يُقلل من جودة التوصية من منظور المستخدم، حيث يُفضل غالبًا اقتراح منتجات مشابهة في الاستخدام والوظيفة، وليس فقط في السعر. على سبيل المثال، قد لا يكون من المنطقي اقتراح "سماعة رأس باهظة الثمن" لمستخدم يتصفح "شاحنًا رخيصًا"، حتى لو كانا في نفس النطاق السعري.

لذلك، يُوصى بعدم إدراج السعر كمُدخل مباشر في نماذج التجميع إلا بعد تقييم دقيق، أو تحويله إلى متغير فئوي (Post-filtering) (مثلًا: "رخيص"، "متوسط"، "غالي") لتقليل تأثيره العددي المفرط. كما يمكن دمج عامل تصفية لاحق (Post-filtering) بدلًا من استخدامه في التجميع الأساسي.

6. الاستنتاج

تُظهر هذه الدراسة أن تضمّن السعر كمُدخل في خوارزميات التوصية ليس دائمًا مفيدًا، وقد يؤدي إلى توصيات غير منطقية من حيث الوظيفة أو الاستخدام. فبينما يُحسّن السعر التماسك المالي داخل العناقيد، فإنه يُضعف التماسك الوظيفي، مما يقلل من جودة التوصية من منظور المستخدم.

التوصيات:

1. تجريب النماذج مع وبدون السعر لتقييم تأثيره على جودة التوصية.
2. تحويل السعر إلى متغير فئوي أو استخدامه كعامل تصفية لاحق بدلًا من تضمينه في التجميع.
3. مع نتائج التجميع لتعزيز دقة التوصيات (Apriori مثل) دمج قواعد الارتباط.
4. الاعتماد على تقييم بشري كمقياس مكمل للمقاييس الكمية.

الأعمال المستقبلية:

(Reinforcement Learning) في المستقبل، يمكن توسيع هذه الدراسة باستخدام بيانات أكبر، أو دمج تعلم التعزيز لتعديل تأثير السعر ديناميكيًا حسب سلوك المستخدم.

7. المراجع

- Ricci, F., Rokach, L., & Shapira, B. (2015). *Recommender Systems Handbook*. Springer.

- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Elsevier.
 - Aggarwal, C. C. (2016). *Recommender Systems: The Textbook*. Springer.
 - Raschka, S. (2023). *mlxtend documentation*. <https://rasbt.github.io/mlxtend/>
 - scikit-learn developers. (2023). *User Guide: Clustering*. <https://scikit-learn.org/>
-

ملاحظة: هذا البحث متاح للاستخدام الأكاديمي والتعليمي. لأي استفسارات أو تطوير إضافي، يُرجى التواصل مع الباحث.