

استخدام تقنيات التنقيب عن البيانات مع نموذج RFM لتقدير القيمة الدائمة للزبون: دراسة حالة

Using Data Mining Techniques with RFM Model to Estimate Customer Lifetime Value: a Case Study

صلاح الدين قوت^{1*}، طارق بلحاج²

¹ المركز الجامعي عبد الحفيظ بوالصوف - ميله، مخبر دراسات استراتيجيات التنوع الاقتصادي لتحقيق التنمية المستدامة (الجزائر)، s.kout@centre-univ-mila.dz

² المركز الجامعي عبد الحفيظ بوالصوف - ميله، مخبر دراسات استراتيجيات التنوع الاقتصادي لتحقيق التنمية المستدامة (الجزائر)، t.belhadj@centre-univ-mila.dz

النشر: 2023/12/31

القبول: 2023/12/31

الاستلام: 2023/09/29

ملخص:

نهدف من خلال هذه الورقة البحثية إلى إيجاد طريقة عملية وبسيطة لتجزئة الزبائن على أساس ربحيتهم وتقدير قيمتهم الدائمة، حيث قمنا بشرح مفهوم القيمة الدائمة للزبون (CLV)، نموذج RFM والتنقيب عن البيانات (DM)، ومن خلال تطبيق نموذج RFM على قاعدة بيانات خاصة بمبيعات مؤسسة صناعية ودمجها مع تقنية من تقنيات التنقيب عن البيانات والمتمثلة في الخوارزمية التجميعية K-means، قمنا بتقدير القيمة الدائمة لكل مجموعة من مجموعات الزبائن المحصل عليها، حيث أظهرت النتائج أن استخدام هذه الطريقة يسهل في تخصيص الموارد ووضع الاستراتيجيات التسويقية المناسبة.

الكلمات المفتاحية: القيمة الدائمة للزبون، نموذج RFM، التنقيب عن البيانات، الخوارزمية التجميعية K-means، إدارة علاقات الزبائن.

Abstract:

Through this paper, we aim to find a simple and practical way to segment customers on the basis of their profitability, as we explain the concept of customer lifetime value (CLV) RFM model and data mining (DM), and by applying the RFM model to an industrial organization transaction database with a data mining technique represented by the aggregate K-means Clustering algorithm, we estimated the lifetime value of each group of customers obtained, the results showed that using this method facilitates the allocation of resources and the development of appropriate marketing strategies.

Keywords: Customer Lifetime Value, RFM Model, Data Mining, K-means Clustering, Customer Relationship Management.

1. مقدمة:

من خلال ما سبق نطرح التساؤل التالي:

- كيف يمكن استخدام تقنيات التنقيب عن البيانات مع نموذج RFM لحساب القيمة الدائمة للزبون؟

سوف نحاول في هذه الدراسة التعريف بالمفاهيم المذكورة مسبقاً، وتطبيق نموذج RFM المدمج مع تقنيات التنقيب عن البيانات على قاعدة بيانات مؤسسة اقتصادية من أجل وضع الزبائن في مجموعات وتقدير القيمة الدائمة لكل مجموعة.

2. مراجعة الأدبيات النظرية:

1.2. القيمة الدائمة للزبون Customer Lifetime Value:

1.1.2. مفهوم القيمة الدائمة للزبون:

يشير كل من (Gupta and Lehmann 2006) أن هناك جانبين من قيمة الزبون يجب فحصهما، القيمة التي تقدمها المؤسسة لزيائنها (وجهة نظر التسويق التقليدية) والقيمة التي يقدمها الزبائن للمؤسسة (وجهة النظر المالية) بطريقة ما يلتقط هذان الجانبان الفلسفات المختلفة للتسويق (الزبون هو الملك) والتمويل (النقد هو الملك)، حيث تحتاج المؤسسة التي تخصص مواردها إلى مراعاة كلا الجانبين لقيمة الزبون، يمكن اعتبار الزبائن الذين لا يقدمون قيمة كبيرة للمؤسسة (قيمة دائمة منخفضة) وفي نفس الوقت يحصلون على قيمة كبيرة من المؤسسة زبائن ذوي ربحية منخفضة، أي أنهم يحصلون على فائدة كبيرة دون تقديم الكثير في المقابل، في بعض الحالات قد يكون من الأفضل للمؤسسة إما رفع الأسعار لهؤلاء

إن التغيرات الحاصلة في السوق مثل تنوع المنتجات وقصر دورة حياتها إضافة إلى ارتفاع توقعات الزبائن، دفعت المؤسسات إلى الانتقال من المفهوم التقليدي للتسويق (التسويق بالصفقات) إلى مفهوم جديد يركز على الزبون (التسويق بالعلاقات) (Hoekstra & Huizingh, 1999)، لقد ساهم هذا المفهوم في ظهور ما يعرف بإدارة علاقات الزبائن (Customer Relationship Management) والتي تهدف في الأساس إلى تحديد الزبائن ذوي القيمة وبناء علاقة طويلة الأمد معهم (Wilson, Daniel, & McDonald, 2002)، حيث أن تحديد الشرائح السوقية المستهدفة يتطلب تجزئة الزبائن إلى مجموعات تحمل خصائص متشابهة (Smith, 1956)، وتعد القيمة الدائمة للزبون الطريقة الأكثر شيوعاً لتقييم ربحية الزبائن والتي تعرف على أنها الإيرادات المتوقعة الحصول عليها من الزبون في المستقبل مطروح منها تكاليف جذبته وخدمته والمحافظة عليه (Heitz, Dettling, & Ruckstuhl, 2011)، من جهة أخرى يعد نموذج RFM أداة بسيطة وسهلة الاستخدام تستخدم بشكل خاص لتحليل سلوك شراء الزبون والذي يتكون من ثلاث متغيرات وهي الحدائة، التكرار والنقدية (Mesforoush & Tarokh, 2013). ولقد أتاحت تكنولوجيا المعلومات استخدام تقنيات التنقيب عن البيانات وإدخالها في مجال التسويق، حيث ظهرت في السنوات الأخيرة العديد من الدراسات التي تدمج تقنيات التنقيب عن البيانات مع نموذج RFM لتقدير القيمة الدائمة للزبون (Shen & Chuang, 2009).

يمكن أن يستهلكه إذا حصل على معاملة خاصة، أما الزبون العادي فسيستلحق امتيازات لا يستحقها.

وفقا لكل من (Estrella-Ramón, 2013) فإن Sánchez-Pérez, Swinnen, and VanHoof أول من تناول مفهوم القيمة الدائمة للزبون (CLV) هو فيليب كوتلر وذلك سنة 1974، حيث عرفها (Kotler 1974, p. 24) تحت مسمى ربحية الزبون على المدى الطويل (long-run Customer Profitability) على أنها "القيمة الحالية لتدفق الربح المستقبلي المتوقع الحصول عليه من الزبون خلال أفق زمني معين"، لكن أول من جاء بهذه الفكرة هو Bursk (1966) تحت مسمى قيمة الاستثمار (Investment Value)، حيث أشار إلى أن الزبائن هم عبارة عن أصول يمكن للمؤسسة الاستثمار فيها.

يعرف (Dwyer 1997, p. 7) القيمة الدائمة للزبون على أنها "القيمة الحالية للمنافع المتوقعة من الزبائن مطروحا منها الأعباء". في حين يعرف كل من (Hoekstra 1999, p. 266) and Huizingh "القيمة الإجمالية للمساهمات المباشرة (الصفقات) والمساهمات غير المباشرة (الكلمة المنطوقة...)) للنفقات العامة والأرباح للزبون الفرد خلال دورة حياته بأكملها، أي من بداية العلاقة حتى نهايتها المتوقعة". كما يعرف كل من (Dorrington and Goodwin 2002, p. 8) القيمة الدائمة للزبون على أنها "قيمة الإيرادات المستقبلية من الزبون، بناء على مقتنياته من المنتجات الحالية والمستقبلية". أما Kumar (2008, p. 6) فيعرف القيمة الدائمة للزبون

الزبائن أو تخفيض قيمة الموارد المخصصة لهم، وفي بعض الأحيان يستحسن التخلي عنهم، بعبارة أخرى قد لا تكون الحصة السوقية الكبيرة مفيدة للمؤسسة وإنما يجب التركيز على نوعية الزبون وحسب كل من (Malthouse and Blattberg 2005) فإن القيمة الدائمة للزبون والتي تمثل القيمة الحالية للمنافع المتوقعة ناقص الأعباء أصبحت موضوعا هاما في التسويق بالعلاقات، حيث يتم تفسير العلاقات مع الزبائن الفرديين على أنها أصول رأسمالية تتطلب إدارة واستثمارا مناسبين بشكل يسمح بتعظيم أرباح المؤسسة من خلال تقييم ربحية الزبائن ثم تصميم برامج تسويقية خاصة بالشرائح السوقية ذات الربحية المرتفعة، قد تبدو هذه الإستراتيجية منطقية بشكل واضح حيث أنه من الشائع أن تمثل نسبة صغيرة من الزبائن نسبة كبيرة من الإيرادات والأرباح، لكن تقييم ربحية الزبائن المستقبل على أساس ربحيتهم الحالية تبقى غير دقيقة، فإذا قدمت المؤسسة لأفضل 20٪ من الزبائن معاملة خاصة فإنها قد تخطئ في تصنيفهم، حيث أنه من بين أفضل 20٪ الفعليين من الممكن أن يتم تصنيف حوالي 55٪ بشكل خاطئ (ولن يتلقوا معاملة خاصة)، كما أنه من بين 80٪ من الزبائن والذين يقعون في الحد الأدنى فعليا 15٪ قد يتم تصنيفهم بشكل خاطئ (وسيتلقون معاملة خاصة)، يؤدي سوء تصنيف الزبائن إلى تكاليف محتملة، يمكن للزبون الأفضل الذي يتم تصنيفه بشكل خاطئ على أنه عادي أن يتحول إلى مؤسسات منافسة أو تطور صورة ذهنية سلبية تجاه المؤسسة، أو لا يستهلك بنفس الشكل الذي

2.1.2. أهمية القيمة الدائمة للزبون:

حسب كل من (Chang, Chang, and (2012) فإن حساب القيمة الدائمة للزبون فائدة ذات شقين: فهم القيمة المحتملة للزبائن، وحث المؤسسات على معرفة المزيد عن أنماط الأفراد أو مجموعات الزبائن، تتيح هذه المعلومات للمؤسسة وضع استراتيجيات مثالية لكل زبون والقضاء على التكاليف المهدرة، وإنشاء منظور طويل الأجل للعلاقة المحتملة مع الزبائن.

وحسب (Kumar (2008 فإنه يمكن فهم أهمية القيمة الدائمة للزبون من خلال تأثيرها على المسألتين التاليتين:

- يساعد حساب القيمة الدائمة للزبون المؤسسة على معرفة مقدار ما يمكن أن تستثمره في الاحتفاظ بالزبون لتحقيق عائد استثمار إيجابي، خاصة وأن موارد المؤسسة محدودة وبالتالي فهي تريد بشكل مثالي الاستثمار في هؤلاء الزبائن الذين يحققون أقصى عائد لها.

- بمجرد أن تقوم المؤسسة بحساب القيم الدائمة لزبائنها يمكنها تخصيص مواردها المحدودة على النحو الأمثل لتحقيق أقصى عائد، كما يعتبر إطار عمل القيمة الدائمة للزبون أيضا أساسا لاختيار الزبائن واتخاذ القرار بشأن استراتيجيات الاتصال الخاصة بالزبائن، حيث يمكن اعتبار القيمة الدائمة للزبون مقياسا يسمح بتخصيص الموارد لأنشطة التسويق المستمرة في المؤسسات التي تتبنى نهجا يركز على الزبون.

على أنها "إجمالي المساهمة المالية من الفترة الحالية إلى المستقبل أي الإيرادات مطروحا منها تكاليف الزبون على مدى حياته/حياتها المستقبلية مع المؤسسة، وبالتالي يعكس الربحية المستقبلية للزبون".

يلاحظ أن التعريف الذي جاء به (1997) Dwyer يأخذ بعين الاعتبار تغير قيمة النقود في المستقبل، كما أنه استخدم لفض "المنافع" بدلا من لفض "الإيرادات"، حيث أن المنافع يمكن أن تكون مباشرة أو غير مباشرة، وهو ما يتوافق مع التعريف الخاص بكل من (1999) Hoekstra and Huizingh والذي يتميز كذلك بأنه يحسب قيمة الزبون من بداية علاقته مع المؤسسة إلى نهايتها، أي مجموع القيمة السابقة والمستقبلية للزبون، وفي حين تتجاهل التعريفات السابقة التكاليف المرتبطة بخدمة الزبون فإن التعريف الذي جاء به (Kumar (2008 يأخذ هذه التكاليف بعين الاعتبار، أما التعريف الذي جاء به كل من (2002) Dorrington and Goodwin فهو الأقرب لدراستنا كون نموذج RFM يتنبأ بربحية الزبون المستقبلية بناء على ربحيته الحالية.

في حين تعرف حقوق ملكية الزبون (Customer Equity) على أنها مجموع القيم الدائمة لجميع زبائن المؤسسة حيث أنه في بعض الحالات وفي حالة عدم وجود نمذجة على المستوى الفردي يتم حساب المقياس الإجمالي لحقوق الزبون مباشرة ومن ثم حساب متوسط القيمة الدائمة للزبون بقسمة حقوق الملكية على عدد الزبائن في قاعدة بيانات المؤسسة (Kumar, Ramani, & Bohling, 2004).

3.1.2. سياقات القيمة الدائمة للزبون ونماذجها:

و يشير كل من (Fader and Hardie (2009) إلى أن السمة الأساسية المحددة للسياق التعاقدى (Contractual Setting) هي أن المؤسسة قادرة على ملاحظة انشقاق الزبون وعادة ما تكون الأسئلة التي تهم الإدارة هي: "أي من الزبائن سيكون لديهم مخاطر أكبر في إحداث انشقاق في الفترة التالية؟" أو "كم من الوقت يمكننا توقع بقاء هذا الزبون معنا، مع إعطاء معلومات حول علاقته بنا حتى الآن؟"، حيث أن الإجابة على هذه الأسئلة تكمن من خلال تصميم أي محاولة لحساب القيمة الدائمة للزبون في السياق التعاقدى، وحسب (Singh and Jain (2017) فإنه من بين النماذج المستخدمة في هذا السياق النموذج الهيكلي لحساب القيمة الدائمة للزبون، والتعبير الرياضي هو:

$$CLV = \sum_{i=1}^n \frac{(R_i - C_i)}{(1 + d)^{i-0,5}}$$

في المعادلة، i هي الفترة التي يجلب فيها الزبائن التدفق النقدي؛ R_i هي إيرادات المؤسسة التي جلبها الزبائن في الفترة t ؛ C_i هي التكلفة الإجمالية المتكبدة لتوليد الإيرادات R_i ؛ $R_i - C_i$ هو ربح هؤلاء الزبائن الذين تم اختبارهم خلال فترة العلاقة؛ n هي دورة حياة الزبون؛ d هو معدل الخصم.

2.2. نموذج RFM:

1.2.2. مفهوم نموذج RFM:

تم اقتراح نموذج RFM لأول مرة بواسطة Hughes من معهد قاعدة البيانات الأمريكي

من بين جميع السياقات التي تمت دراستها يعتبر قياس القيمة الدائمة للزبون أكثر تعقيدا في السياق غير التعاقدى (Noncontractual Setting)، حيث أن السياق غير التعاقدى هو السياق الذي لا تلاحظ فيه المؤسسة انشقاق الزبون¹ (Customer Churn) ويعد هذا السياق الأكثر شيوعا في الممارسة، ومن أمثلة هذا السياق مشتريات التجزئة حيث لا تعرف المؤسسة ما إذا كان الزبون قد انشق أم لا ولقياس القيمة الدائمة للزبون في السياقات غير التعاقدية يحتاج المرء أولا إلى التنبؤ بمحركات القيمة الدائمة للزبون، أي عمر الزبون ومشترياته، وهناك نموذجان معترف بهما على نطاق واسع لهذا الغرض هما نموذج Pareto/NBD والذي يسمح بتقدير عدد الزبائن النشطين حاليا، كيف يتغير هذا العدد بمرور الوقت، احتمالية نشاط زبون معين، المدة التي من المحتمل أن يبقى فيها الزبون نشطا، إضافة إلى العدد المتوقع لعمليات الشراء التي من الممكن أن يقوم بها الزبون خلال فترة زمنية مستقبلية، أما النموذج الثاني فهو نموذج BG/NBD حيث أن هذا الأخير هو عبارة عن

¹ يعرف كل من (Gady, Baesens, and (2009, p. 403) Croux انشقاق الزبون (Customer Churn) على أنه "الحالة التي يتوقف فيها الزبون عن التعامل مع المؤسسة، كما أن الزبون المهتد بالانشقاق هو الزبون الذي تكون قيمته الدائمة في تناقص مع مرور الوقت، في هذه الحالة لا يزال لديه علاقة مع المؤسسة لكنه سوف يتحول إلى منافس في المستقبل القريب".

درجات RFM معا، حيث يحصل أفضل الزبائن على درجة مركبة 15 أي (5+5+5) ويحصل أسوأ الزبائن على درجة مركبة 3 أي (1+1+1) كحد أدنى، نظرا لأن هذه المتغيرات معروفة على نطاق واسع فغالبا ما يتم تضمينها في نماذج القيمة الدائمة للزبون (Blattberg, Malthouse, & Neslin, 2009) إضافة إلى طريقة التنقيب قام العديد من الباحثين بدمج نموذج RFM مع أدوات أخرى، حيث قام كل من Colombo and Jiang (1999) بتطوير نموذج RFM عشوائي (Stochastic RFM Model) والذي يقدم مزايا من بينها سهولة التحديث والقدرة على إضافة متغيرات أخرى كالمتغيرات الديموغرافية مثلا، ومن أجل تحديد الأهمية النسبية لأبعاد نموذج RFM اقترح كل من Shih & Liu (2003) عملية التحليل الهرمي (Analytic Hierarchy Process) . مع استخدام خوارزمية K-means التجميعية من أجل وضع الزبائن في مجموعات تحمل خصائص متشابهة من حيث أبعاد النموذج، وجاء كل من Fader, Hardie, and Lee (2005) بفكرة منحنيات القيمة المتساوية (Iso-Value Curves) والتي تتيح تجميع الزبائن الأفراد الذين لديهم تاريخ شراء مختلف وتصور التفاعلات والمفاضلات بين قياسات نموذج RFM والقيمة الدائمة للزبون، حيث يعتمد النموذج العشوائي على إطار عمل نموذج Pareto/NBD للتقاط تدفق المعاملات بمرور الوقت ونموذج فرعي Gamma-Gamma للإنفاق لكل معاملة، استخدم كل من Chen, Kuo, Wu, and Tang (2009) طريقة استخراج الأنماط المتعددة (Sequential Pattern Mining) ودمجها مع

سنة 1994 كأداة شائعة لتحليل قيمة الزبون فقد تم استخدامه على نطاق واسع في تجزئة الزبائن وتحليل سلوكهم من خلال التنقيب في قواعد بيانات المبيعات، RFM هو اختصار للحدثة (Recency) التكرار (Frequency) والنقدية (Monetary)، والتي تشير إلى حدثة آخر عملية شراء وتكرار الشراء والقيمة النقدية للشراء على التوالي، تشير الحدثة إلى الزمن المنقضي منذ آخر عملية شراء قام بها الزبون، يشير التكرار إلى عدد المرات التي قام فيها الزبون بعملية شراء خلال فترة زمنية معينة، في حين تمثل النقدية المبلغ الإجمالي الذي ينفقه الزبون في الشراء خلال الفترة الإحصائية، تجدر الإشارة إلى أنه كلما قل الزمن المنقضي منذ آخر عملية شراء وزاد التكرار والنقدية كلما كان ذلك مؤشرا على ارتفاع قيمة الزبون (Cheng & Chen, 2009).

2.2.2. طرق تحليل ربحية الزبون وفق نموذج RFM:

تعد طريقة التنقيب (Scoring Method) الطريقة الأساسية لنموذج RFM، حيث يتم استخراج قيم RFM للزبائن المؤسسة من قاعدة بيانات مبيعاتها، ثم يتم تقسيم قاعدة البيانات إلى خمسة أقسام متساوية وإعطاء أرقام من 1 إلى 5 لكل من المتغيرات الثلاثة مما ينتج عنه (5x5x5) أو 125 مجموعة (خلية) من الزبائن حيث أن الزبون الذي ينتهي إلى أعلى 20% من الزبائن من حيث الحدثة والتكرار والنقدية سوف يحصل على تقييم 555، أي أنه سوف يكون في المجموعة الأعلى تقييما من حيث الربحية (McCarty & Hastak, 2007). بدلا من ذلك يفضل Miglautsch (2000) إضافة

3.2.2. تطور متغيرات نموذج RFM:

قام الباحثون بإضافة متغيرات جديدة حيث جاء كل من Yeh, Yang, and Ting (2009) بنموذج RFMTC حيث يشير T إلى الزمن المنقضي من أول عملية شراء (time since first purchase)، ويشير C إلى احتمال فقدان الزبون (Churn Probability) ومن خلال استخدام تسلسل برنولي (Bernoulli Sequence) في نظرية الاحتمالات يتم استخراج الصيغ التي يمكن أن تتنبأ باحتمالية أن يشتري زبون واحد في المرة القادمة والقيمة المتوقعة لإجمالي عدد المرات التي سيشتري فيها الزبون في المستقبل، وجاء كل من Chang and Tsai (2011) بنموذج GRFM حيث يشير G إلى المجموعة (Group) ويقصد بها مجموعة العناصر المشتراة، حيث أن قيمة الزبون تختلف حسب العناصر التي يقوم بشرائها، أي أنه قد ينتهي إلى أكثر من مجموعة حسب العناصر المشتراة. وجاء كل من Zhang, Bradlow, and Small (2015) بنموذج RFMC حيث يشير C إلى التكتل (Clumpiness) وهو درجة عدم المطابقة للمسافات المتساوية ويظهر خصائصه كمكون مهم في فهم أفضل للقيمة الدائمة للزبون ومكوناتها الفرعية. وجاء كل من Heldt, Silveira, and Luce (2021) بنموذج RFM/P، حيث يشير P إلى المنتج (Product)، مما يسمح بتقدير القيمة الدائمة للزبون بشكل أفضل تبعاً لمحفظة مشترياته. في حين جاء كل من Ozkan and Deveci (2021) بنموذج RFM-V، حيث يشير V إلى التنوع (Variety) أي تنوع مشتريات الزبون، والتي يمكن أن تحدد عمق الزبون من حيث أدبيات إدارة علاقات الزبائن، كما تقترح

نموذج RFM مما يسمح بتقديم إطار عمل لتوليد معلومات قيمة حول أنماط شراء الزبائن وتغيرها مع مرور الوقت، كما قام كل من Cheng and Chen (2009) بتجزئة الزبائن وفق نموذج RFM والخوارزمية التجميعية K-means، إضافة إلى الاعتماد على نظرية المجموعة التقريبية (Rough Set Theory) والتي تعتمد على النمذجة الرياضية للتعامل مع مشاكل تصنيف بيانات الفئة، ثم تضمين هذه البيانات في عملية اتخاذ القرار، وهذا ما يقود إلى إدارة علاقات زبائن ممتازة، بدلا من ذلك استخدم كل من Safari, Safari, and (2016) خوارزمية fuzzy C-means التجميعية والتي تأخذ بعين الاعتبار إمكانية انتماء الزبون لأكثر من مجموعة واحدة وعملية التحليل الهرمي الضبابية (Fuzzy AHP Method) كبديلين للخوارزمية التجميعية K-means وعملية التحليل الهرمي العادية، استخدم كل من Azadnia, Saman, (2011) Wong, and Hemdi الخوارزمية التجميعية fuzzy-C-means مع الاعتماد على تقنية تفضيل الطلب عن طريق التشابه مع الحل المثالي (Technique for Order Preference by Similarity to Ideal Solution) من أجل ترتيب الزبائن على أساس قيمهم الدائمة، كما قام كل من Berahmana, Mohammed, and (2020) Chairuang بتجزئة الزبائن وفق نموذج RFM مع اختبار ثلاثة أنواع من الخوارزميات التجميعية والمتمثلة في K-means، K-medoids و DBSCAN، حيث أظهرت النتائج أن K-means لديها أحسن مستوى من الصلاحية وتعطي نتائج أكثر موثوقية.

الدراسة أيضا مصفوفة جديدة تسمى بمصفوفة عمق المنتج للزبون، والتي تسمح بتطوير استراتيجيات تسويق أكثر صحة وتحقيق اقتراحات ترويجية فعالة.

حسب كل من Fayyad, Piatetsky- (1996) Shapiro, and Smyth فإن التنقيب عن البيانات هو خطوة معينة من الخطوات التي تندرج تحت عملية اكتشاف المعرفة في قواعد البيانات (Knowledge Discovery in Databases)، والتي تتكون من تطبيق تحليل البيانات وخوارزميات الاكتشاف التي تنتج تعدادا معيناً للأنماط (أو النماذج) على البيانات، حيث يشير كل من Zhang, (2003) Zhang, and Yang إلى أن عملية اكتشاف المعرفة في قواعد البيانات تتكون من أربع خطوات رئيسية وهي:

- تحديد المشكلة (Defining the Problem): تحديد الأهداف من اكتشاف المعرفة وقابليتها للتنفيذ، إضافة إلى تحديد البيانات التي سوف يتم استخدامها.
- تحضير البيانات (Data Pre-processing): وتضم جمع البيانات، تكامل البيانات تحويل البيانات، تنقية البيانات، تقليل البيانات وتقدير البيانات.
- التنقيب عن البيانات (Data Mining): يتم تطبيق طرق ذكية لاستخراج أنماط البيانات.
- ما بعد التنقيب عن البيانات (Post Data Mining): يتم تقييم الأنماط، نشر النموذج، الصيانة وعرض المعرفة.

التنقيب عن البيانات هو مجال بحث وتطبيق متعدد التخصصات يهدف إلى اكتشاف معرفة جديدة ومفيدة من قواعد البيانات الضخمة، من خلال استخدام أساليب تتراوح بين الذكاء الاصطناعي والإحصاءات وقواعد البيانات، تم استخدام تقنيات التنقيب عن البيانات في المجالات التي تحتوي على بيانات

الدراسة أيضا مصفوفة جديدة تسمى بمصفوفة عمق المنتج للزبون، والتي تسمح بتطوير استراتيجيات تسويق أكثر صحة وتحقيق اقتراحات ترويجية فعالة.

4.2.2. إيجابيات وسلبيات نموذج RFM:

لنموذج RFM العديد من الإيجابيات من بينها سهولة الاستخدام والفعالية في تقدير القيمة الدائمة للزبون (Customer Lifetime Value)، يسمح بتجزئة الزبائن على أساس ربحيتهم، إضافة إلى زيادة الأرباح على المدى القصير، كما لا يخلو هذا النموذج من العيوب ومن بينها الاعتماد على ثلاثة أبعاد فقط وتجاهل العديد من السمات مثل الدخل وأسلوب الحياة وتنوع المنتجات، كما أنه يعتمد على معطيات مسبقية (سجلات الشراء) لقياس ربحية الزبائن الحاليين وبالتالي فهو غير قادر على قياس ربحية الزبائن المحتملين، إضافة إلى اختلاف الأهمية النسبية لأبعاد النموذج الثلاثة من صناعة إلى أخرى (Dursun & Caber, 2016)، كما يعتبر Kumar (2008) أن نموذج RFM هو أداة تقليدية لإدارة ولاء الزبون، يتمثل أحد أوجه القصور الرئيسية في كونه ليس مقياسا تطلعيا وبالتالي لا يتنبأ فيما إذا كان الزبون سيكون نشطا في المستقبل، فهو يأخذ فقط في الاعتبار سلوك الشراء الملحوظ ويستنبطه في المستقبل للوصول إلى الربحية المستقبلية للزبون، خاصة وأن الأهمية النسبية لأبعاده تؤثر بشكل كبير على قياس ربحية الزبون.

3.2. التنقيب في البيانات Data Mining:

1.3.2. مفهوم التنقيب عن البيانات:

نفس الوقت ويمكن أن تكون الخوارزميات الهرمية تكتلية (من أسفل إلى أعلى) أو تقسيمية (من أعلى إلى أسفل)، حيث تبدأ الخوارزميات التكتلية بكل عنصر كمجموعة منفصلة وتدمجها في مجموعات أكبر متتالية، بينما تبدأ الخوارزميات الانقسامية بالمجموعة بأكملها وتنتقل إلى تقسيمها إلى مجموعات أصغر متتالية.

3.3.2. خوارزمية K-means التجميعية:

تعد خوارزمية K-means واحدة من أكثر الخوارزميات التجميعية (خوارزميات التحليل العنقودي) شيوعاً، استخدمت لأول مرة بواسطة J.B MacQueen في عام 1967، يشير الحرف K في اسمه إلى حقيقة أن الخوارزمية تبحث عن عدد ثابت من المجموعات التي يتم تحديدها من حيث قرب نقاط البيانات من بعضها البعض، وعادة ما تعامل الخوارزمية مع أكثر من متغيرين مستقلين (Mesforoush & Tarokh, 2013). تبدأ خوارزمية K-means بعدد K من النقاط الوسطى (Centroids) (يتم اختيار القيم الأولية للنقاط الوسطى بشكل عشوائي أو اشتقاقها من معلومات مسبقة)، بعد ذلك يتم تعيين كل نمط في مجموعة البيانات إلى أقرب مجموعة (أي أقرب نقطة وسطى)، أخيراً يتم إعادة حساب النقاط الوسطى وفقاً للأنماط المرتبطة بها حيث تتكرر هذه العملية حتى يتحقق التقارب (Madhulatha, 2012). تم استخدام هذه الخوارزمية على نطاق واسع في مختلف المجالات بما في ذلك استخراج وتحليل البيانات الإحصائية وتطبيقات الأعمال الأخرى ومن بين أهم تطبيقات خوارزمية K-means في مجال

منظمة مثل إدارة علاقات الزبائن في الخدمات المصرفية والتجزئة تركز هذه التقنيات على الوصول إلى معرفة جديدة ومفيدة من خلال استكشاف الأنماط التي تكون مخفية داخل البيانات الضخمة (Big Data). توفر هذه الأنماط معلومات مفيدة للمؤسسات حول اتجاهات الأسواق والزبائن والموردين والمنافسين وغير ذلك (Shen & Chuang, 2009).

2.3.2. طرق التنقيب عن البيانات:

هناك طريقتان للتنقيب عن البيانات: وصفية وتنبؤية، حيث أن التصنيف (Classification) هو طريقة تنبؤية والتحليل العنقودي (Cluster Analysis) هو طريقة وصفية، ويعرف التصنيف على أنه عملية العثور على نموذج (أو وظيفة) تسمح بوصف وتميز فئات أو مفاهيم البيانات، بغرض التمكن من استخدام النموذج للتنبؤ بفئة الكائنات التي يكون تصنيف فئتها غير معروف (Khajvand, 2011). أما الهدف من التحليل العنقودي هو تصنيف الوحدات (سواء كانت أفراداً أو أشياء) بحيث يكون هناك تشابه أكبر بين الوحدات داخل المجموعات أكثر من بين الوحدات في مجموعات مختلفة (Klasterin, 1983). وأما التحليل العنقودي (التجميع) فيضم طريقتين وهما التجميع الهرمي (Hierarchical Clustering) والتجميع الجزئي (Partitional Clustering) حيث تجد الخوارزميات الهرمية عناقيد (مجموعات) متتالية باستخدام مجموعات منشأة مسبقاً، بينما تحدد الخوارزميات الجزئية جميع المجموعات في

وتقدير القيمة الدائمة لمجموع الزبائن وفق الخطوات التالية:

- تجهيز ومعالجة البيانات.
- استخراج قيم RFM إضافة إلى التنقيط الخاص بكل زبون.
- تطبيق خوارزمية K-means التجميعية على قيم RFM، ومقارنتها مع هرم الزبائن.
- تنميط قيم RFM الخاصة بكل مجموعة.
- تقدير القيمة الدائمة الخاصة بكل مجموعة.

2.3. النتائج:

يعرض الجدول رقم 1 عينة لقاعدة البيانات محل الدراسة.

جدول 1: عينة لقاعدة البيانات محل الدراسة

رقم الزبون	تاريخ الشراء	المبلغ (دج)
280048	02 جانفي 2021	2 500 000,00
280081	02 جانفي 2021	2 000 000,00
190203	02 جانفي 2021	754 959,00
....
....
310087	30 ديسمبر 2021	1 500 000,00

المصدر: قاعدة بيانات المؤسسة

بعد تنظيم قاعدة البيانات يتم إدخالها في برنامج IBM SPSS statistics 24 والذي يتيح إجراء تحليل RFM بشكل سهل وبسيط. مع اعتماد تاريخ 01 جانفي 2022 ليكون تاريخ إجراء الحساب، فمثلا الزبون الذي يحمل الرقم 040133 قام بأخر عملية شراء بتاريخ 24 ديسمبر 2021 ما يجعله من بين أعلى 20% من الزبائن من حيث الحداثة، وهو ما يفسر حصوله على تقييم 5، أما قيامه بعملية شراء

التسويق تجزئة الزبائن، حيث يتم استخدامها مع أبعاد نموذج RFM من أجل تجزئة الزبائن إلى مجموعات يحمل أفرادها خصائص متشابهة مما يسمح بالتنبؤ بالقيمة الدائمة للزبائن كل مجموعة (Wu, et al., 2020).

ويشير كل من Omran, Engelbrecht, (2007) and Salman إلى أن خوارزمية K-means تتمتع بمجموعة من المزايا من بينها سهولة التطبيق ومناسبتها لعدد كبير من البيانات، غير أنها لا تخلو من العيوب ومن بينها ما يلي:

- اعتمادها على البيانات (Data-dependent).
- إنها خوارزمية جشعة (Greedy Algorithm)، حيث تعتمد على الظروف الأولية مما قد يتسبب في تقارب الخوارزمية مع حلول دون المستوى الأمثل.
- تحتاج إلى تحديد عدد المجموعات بشكل مسبق.

3. دراسة حالة

1.3. المنهجية:

تم إنجاز هذه الدراسة في مؤسسة مختصة في إعادة تحويل الحديد (نصف مصنع) والتجارة بالجملة في الصفائح الحديدية، حيث تم الحصول على قاعدة بيانات مبيعات المؤسسة من خلال برنامج إدارة الموارد الخاص بها تحتوي قاعدة البيانات على اسم ورقم الزبون، تاريخ الشراء والمبلغ الإجمالي للشراء، تم استخدام بيانات سنة 2021 (من 2021/01/02 إلى 2021/12/30) والتي تضم 980 زبونا و2395 عملية شراء، وقد تم استخراج قيم RFM

الخطوة التالية هي وضع الزبائن في مجموعات باستخدام خوارزمية K-means، وحيث أن K يمثل عدد المجموعات فقد تم ضبطه عند الرقم 4 بما يتماشى مع هرم الزبائن الذي جاء به كل من (2001) Zeithaml, Rust, and Lemon والذي يتكون من 4 درجات (البلاتين، الذهب، الحديد والرصاص)، والجدول رقم 3 يظهر نتائج خوارزمية K-means لقبم RFM.

تظهر النتائج أن المجموعة رقم 2 والتي تضم زبونا واحدا هي المجموعة التي تتوافق مع التصنيف البلاتيني في هرم الزبائن حيث أنها تحمل النمط ($R \downarrow F \uparrow M$)، أي أن هذا الزبون لم يمر وقت طويل منذ آخر مرة قام فيها بعملية شراء، كما أنه يتسم بالولاء نظرا لامتلاكه تردد مرتفع، وإنفاقه عالي.

المجموعة رقم 1 والتي تحمل النمط ($R \uparrow M \uparrow F$)، تضم 3 زبائن وهي تتوافق مع التصنيف الذهبي في هرم الزبائن، هذه المجموعة لديها تردد ونقدية جيدين لكن ليس بقدر المجموعة رقم 2، كما أن زبائن هذه المجموعة قد مرت عليهم مدة معتبرة منذ آخر عملية شراء قاموا بها.

المجموعة رقم 4 والتي تحمل النمط ($R \uparrow F \downarrow M$)، تضم 38 زبونا وهي تتوافق مع التصنيف الحديدي في هرم الزبائن هذه المجموعة لديها تردد ونقدية ضعيفين، كما أن زبائن هذه المجموعة قد مرت عليهم مدة طويلة منذ آخر عملية شراء قاموا بها.

واحدة فقط فقد ترجم حصوله على تقييم 2 أي أنه من بين أدنى 40% من الزبائن من حيث تكرار الشراء، في حين أن إنفاقه لمبلغ 1266900.00 دج فقد ترجم حصوله على تقييم 3 أي أنه من بين أعلى 60% من الزبائن من حيث النقدية، والتنقيط النهائي هو 524 أي أن هذا الزبون لديه حدائة مرتفعة وتردد ضعيف ونقدية متوسطة، كما يمكن أن نفسر الحدائة المرتفعة والقيام بعملية شراء واحدة بأن هذا الزبون هو زبون جديد، ونتائج التحليل موضحة في الجدول رقم 2 (تنقيط RFM يكون بنفس ترتيب الأحرف باللغة الأجنبية).

تظهر نتائج تحليل RFM أنه من بين 980 زبونا فقط 09 زبائن تحصلوا على تنقيط 555 ما نسبته 0,1% من الزبائن وقد ساهموا بنسبة 19% من رقم أعمال المؤسسة في سنة 2021، كما أن الزبون الأعلى تقييما قد ساهم لوحده بما نسبته 11% من رقم أعمال المؤسسة، في حين أن 196 زبونا أي ما نسبته 20% من الزبائن قد ساهموا بنسبة 75% في رقم أعمال المؤسسة، هذه النتائج تتوافق بشكل كبير مع قانون باريتو الذي يقول أن 20% من الزبائن يخلقون 80% من الثروة للمؤسسة، وفي نفس الوقت فإن النتائج تعطينا مؤشرات هامة حول تفاوت مساهمة أعلى 20% من الزبائن في رقم أعمال المؤسسة، حيث أن أعلى 13% من زبائن هذه الشريحة قد ساهموا بما نسبته 50% من مساهمة هذه الشريحة، أي أن أعلى 26 زبونا من حيث النقدية ساهموا بنفس القدر الذي ساهم به 170 زبونا الذين يلونهم في الترتيب.

جدول 2: نتائج تحليل RFM

رقم الزبون	تاريخ أحدث شراء	التكرار	النقدية (دج)	تنقيط RFM
040133	24 ديسمبر 2021	1	1 266 900,00	523
040096	01 جويلية 2021	2	2 900 000,00	243
020069	15 ديسمبر 2021	1	1 000 000,00	522
....
....
280094	15 أوت 2021	1	500 000,00	321

المصدر: من إعداد الباحثين بالاعتماد على مخرجات برنامج SPSS

جدول 3: نتائج خوارزمية K-means لقيم RFM

المجموعة	عدد الزبائن	الحدائة R (الأيام)	التكرار F	النقدية M (دج)	النمط
1	3	99,33	19	127 496 653,20	R↑ F↑ M↑
2	1	6,00	35	579 479 983,10	R↓ F↑ M↑
3	938	133,90	2	3 056 911,08	R↑ F↓ M↓
4	38	79,34	11	40 265 822,48	R↑ F↓ M↓
المجموع/المتوسط	980	79,64	16,75	187 574 842,00	-

المصدر: من إعداد الباحثين بالاعتماد على مخرجات برنامج SPSS

المعادلات التي جاء بها كل من (2009) Shen and Chuang والتالية:

$$R' = \frac{R_{max} - R}{R_{max} - R_{min}}$$

$$F' = \frac{F - F_{min}}{R_{max} - R_{min}}$$

$$M' = \frac{M - M_{min}}{R_{max} - R_{min}}$$

حيث تمثل R', F', M' قيم RFM بعد التنميط، $R_{max}, F_{max}, M_{max}$ تمثل أعلى قيم R، F، M و $R_{min}, F_{min}, M_{min}$ تمثل أقل

المجموعة رقم 3 والتي تحمل النمط $(R↑ F↓ M↓)$ ، تضم 938 زبونا وهي تتوافق مع التصنيف الرصاصي في هرم الزبائن هذه المجموعة لديها تردد ونقدية ضعيفين جدا، كما أن زبائن هذه المجموعة قد مرت عليهم مدة معتبرة منذ آخر عملية شراء قاموا بها.

الخطوة التالية هي تقدير القيمة الدائمة لكل مجموعة من الزبائن، ونظرا لأن قيم RFM تختلف من حيث الوحدات (أيام، دينار جزائري...) فإننا سوف نقوم بتنميطها وفق

المجموعة رقم 1 والتي تضم 3 زبائن تحتل المرتبة الثانية من حيث القيمة الدائمة لزبائنها (CLV=0,442)، أي أنه من المتوقع أن يساهموا بنسبة معتبرة في رقم أعمال المؤسسة ويستمرروا في التعامل معها لفترة ليست بالقصيرة، زبائن هذه المجموعة لديهم ربحية وولاء جديدين مع وجود حساسية للأسعار، لذلك فهم عادة ما يتعاملون مع عدة مؤسسات لتقليل المخاطر وبالتالي فإن المؤسسة مطالبة بتكثيف عروضها مع حاجاتهم ورغباتهم مما يسمح بنقلهم إلى مستوى أعلى من الربحية والولاء.

المجموعة رقم 4 والتي تضم 38 زبونا احتلت المرتبة الثالثة من حيث القيمة الدائمة لزبائنها (CLV=0,354)، حيث أنه من المتوقع أن يساهموا بنسبة ضعيفة في رقم أعمال المؤسسة ويستمرروا في التعامل معها لفترة ليست بالقصيرة، زبائن هذه المجموعة لديهم ولاء جيد ولكن ربحيتهم ضعيفة، يجب على المؤسسة دراسة سلوكهم وتحديد العوامل التي تتحكم في مستويات الرضا لديهم، ومن خلال التأثير على هذه العوامل يمكن للمؤسسة تحسين درجة المشاركة في محفظتهم.

المجموعة رقم 3 والتي تضم 938 زبونا احتلت المرتبة الرابعة من حيث القيمة الدائمة لزبائنها (CLV=0,220)، حيث أنه من المتوقع أن يساهموا بنسبة ضعيفة جدا في رقم أعمال المؤسسة، فهم يفتقرون إلى الولاء والربحية وعلاوة على ذلك فإن زبائن هذه المجموعة قد يكلفون المؤسسة الموارد والأموال لذلك يجب إيجاد طرق لتخفيض تكاليف خدمتهم، كما يجب عدم إهمال حقيقة أن هذه المجموعة

قيم R, F, M لكل الزبائن فعلية التنميط سوف تسمح بحصر كل القيم بين 0 و1، وسوف يتم تقدير القيمة الدائمة لكل مجموعة وفق المعادلة التالية:

$$CLV = WR * R' + WF * F' + WM * M'$$

حيث تمثل WR, WF و WM أوزان قيم RFM والتي سوف نفترض أنها متساوية وبالتالي تقدر كما يلي:

$$WR = WF = WM = 1/3 = 0,333$$

الجدول رقم 4 يعرض نتائج تقدير القيمة (CLV) الدائمة للمجموعات الأربعة، وذلك بعد تنميط قيم RFM .

تظهر النتائج أن نتائج تقدير القيمة الدائمة للمجموعات الأربعة تتوافق مع نتائج خوارزمية K-means لقيم RFM حيث حصلت المجموعة رقم 2 على أعلى تقييم، تليها المجموعة الأولى ثم المجموعة الرابعة، في حين حصلت المجموعة الثالثة على أقل تقييم، ويمكن تحليل هذه النتائج كما يلي:

المجموعة رقم 2 والتي تضم زبونا واحدا والذي يحتل المرتبة الأولى من حيث قيمته الدائمة (CLV=0,898)، أي أنه من المتوقع أن يساهم بنسبة كبيرة في رقم أعمال المؤسسة ويستمر بالتعامل معها لفترة طويلة، فهو زبون ذو ربحية مرتفعة وولاء عالي جدا، يتميز بانخفاض حساسيته للأسعار واستعداده لتجربة العروض الجديدة التي تقدمها المؤسسة، وعادة ما يستخدم الكلمة المنطوقة وبالتالي فإن المؤسسة مطالبة بالحفاظ عليه من خلال تزويده بخدمات ذات جودة عالية.

تمثل أكبر شريحة من زبائن المؤسسة وبالتالي فإنه من الوارد جدا أنها تخفي زبائن يمكن تحويلهم إلى زبائن مربحين.

جدول 4: نتائج تقدير القيمة الدائمة للمجموعات الأربعة

ترتيب CLV	CLV	النقدية M' (دج)	التكرار F'	الحدائة R' (الأيام)	المجموعة
2	0,442	0,220	0,375	0,731	1
1	0,898	1,000	0,708	0,989	2
4	0,220	0,005	0,021	0,636	3
3	0,354	0,069	0,208	0,786	4

المصدر: معالجة البيانات

5. قائمة المراجع

1. Azadnia, A. H., Saman, M. Z., Wong, K. Y., & Hemdi, A. R. (2011). Integration model of Fuzzy C means clustering algorithm and TOPSIS Method for Customer Lifetime Value Assessment. 2011 IEEE International Conference on Industrial Engineering and Engineering Management (pp. 16-20). IEEE.
2. Berahmana, R. W., Mohammed, F. A., & Chairuang, K. (2020). Customer segmentation based on RFM model using K-means, K-medoids, and DBSCAN methods. Lontar Komputer: Jurnal Ilmiah Teknologi Informasi, 11 (1), 32-43.
3. Blattberg, R. C., Malthouse, E. C., & Neslin, S. A. (2009). Customer lifetime value: Empirical generalizations and some conceptual questions. Journal of Interactive Marketing, 23 (2), 157-168.

4. خاتمة:

في هذه الدراسة قمنا بتطبيق نموذج RFM على قاعدة بيانات خاصة بمبيعات مؤسسة صناعية، والذي سمح بتحليل سلوك شراء كل زبون، حيث أظهرت نتائج تحليل ربحية الزبائن وفق نموذج RFM أن شريحة صغيرة من الزبائن تساهم بنسبة كبيرة في أرباح المؤسسة، وهو ما دعمته نتائج تحليل النقدية التي توافقت بشكل كبير مع قانون باريتو، ومن خلال استخدام خوارزمية K-means تم وضع الزبائن في مجموعة تحمل أنماط متشابهة لقيم RFM ومقارنتها مع هرم الزبائن مما سمح بتقييم ربحية وولاء كل مجموعة، وبعد تنميط قيم RFM قمنا بتقدير القيمة الدائمة لكل مجموعة، حيث أن هذه التقديرات المتعلقة بالقيمة الدائمة للزبائن توفر لمدرء المؤسسات طريقة عملية لوضع استراتيجياتها التسويقية (اتخاذ القرار التسويقي المناسب)، كما أن هذه النتائج تدعم التوجه نحو مفهوم التسويق

12. Dwyer, F. R. (1997). Customer lifetime valuation to support marketing decision making. *Journal of Direct Marketing* , 11 (4), 6-13.
13. Estrella-Ramón, A., Sánchez-Pérez, M., Swinnen, G., & VanHoof, K. (2013). A marketing view of the customer value: Customer lifetime value and customer equity. *South African Journal of Business Management* , 44 (4), 47-64.
14. Fader, P. S., & Hardie, B. G. (2009). Probability Models for Customer-Base Analysis. *Journal of interactive marketing* , 23 (1), 61-69.
15. Fader, P. S., Hardie, B. G., & Lee, K. L. (2005). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of marketing research* , 42 (4), 415-430.
16. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine* , 17 (3), 37-54.
17. Gladly, N., Baesens, B., & Croux, C. (2009). Modeling churn using customer lifetime value. *European Journal of Operational Research* , 197 (1), 402–411.
18. Gupta, S., & Lehmann, D. R. (2006). Customer Lifetime Value and Firm Valuation. *Journal of Relationship Marketing* , 5 (2-3), 87-110.
19. Heitz, C., Dettling, M., & Ruckstuhl, A. (2011). Modelling customer lifetime value in contractual settings. *International Journal of Services*
4. Bursk, E. C. (1966). View your customers as investments. *Harvard Business Review* , 44 (3), 91-94.
5. Chang, H.-C., & Tsai, H.-P. (2011). Group RFM analysis as a novel framework to discover better customer consumption behavior. *Expert Systems with Applications* , 38 (12), 14499-14513.
6. Chang, W., Chang, C., & Li, Q. (2012). Customer lifetime value: A review. *Social Behavior and Personality: an international journal* , 40 (7), 1057-1064.
7. Chen, Y.-L., Kuo, M.-H., Wu, S.-Y., & Tang, K. (2009). Discovering recency, frequency, and monetary (RFM) sequential patterns from customers' purchasing data. *Electronic Commerce Research and Applications* , 8 (5), 241-251.
8. Cheng, C.-H., & Chen, Y.-S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications* , 36 (3), 4176-4184.
9. Colombo, R., & Jiang, W. (1999). A stochastic RFM model. *Journal of Interactive Marketing* , 13 (3), 2-12.
10. Dorrington, P., & Goodwin, J. (2002). The role of lifetime value in customer relationship management. *Interactive Marketing* , 4 (1), 7-18.
11. Dursun, A., & Caber, M. (2016). Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. *Tourism management perspectives* , 18, 153-160.

28. Malthouse, E. C., & Blattberg, R. C. (2005). Can we predict customer lifetime value? *Journal of interactive marketing* , 19 (1), 2-16.
29. McCarty, J. A., & Hastak, M. (2007). Segmentation approaches in data-mining: A comparison of RFM, CHAID, and logistic regression. *Journal of business research* , 60 (6), 656-662.
30. Mesforoush, A., & Tarokh, M. (2013). Customer profitability segmentation for SMEs case study: network equipment company. *International Journal of Research in Industrial Engineering* , 2 (1), 30-44.
31. Miglautsch, J. R. (2000). Thoughts on RFM scoring. *Journal of Database Marketing & Customer Strategy Management* , 8 (1), 67-72.
32. Omran, M. G., Engelbrecht, A. P., & Salman, A. (2007). An overview of clustering methods. *Intelligent Data Analysis* , 11 (6), 583–605.
33. Ozkan, P., & Deveci Kocakoc, I. (2021). A Customer Segmentation Model Proposal for Retailers: RFM-V. *University of South Florida M3 Center Publishing* , 5 (2021), 1-12.
34. Safari, F., Safari, N., & Montazer, G. A. (2016). Customer lifetime value determination based on RFM model. *Marketing Intelligence & Planning* , 34 (4).
35. Shen, C.-C., & Chuang, H.-M. (2009). A study on the applications of data mining techniques to enhance customer lifetime Technology and Management , 16 (2), 172-190.
20. Heldt, R., Silveira, C. S., & Luce, F. B. (2021). Predicting customer value per product: From RFM to RFM/P. *Journal of Business Research* , 127, 444-453.
21. Hoekstra, J. C., & Huizingh, E. K. (1999). The lifetime value concept in customer-based marketing. *Journal of Market-Focused Management* , 3 (3), 257-274.
22. Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: case study. *Procedia Computer Science* , 3, 57-63.
23. Klastorin, T. (1983). Assessing Cluster Analysis Results. *Journal of Marketing Research* , 20 (1), 92-98.
24. Kotler, P. (1974). Marketing during periods of shortage. *Journal of marketing* , 38 (3), 20-29.
25. Kumar, V. (2008). Customer Lifetime Value—The Path to Profitability. *Foundations and trends® in marketing* , 2 (1), 1-96.
26. Kumar, V., Ramani, G., & Bohling, T. (2004). Customer lifetime value approaches and best practice applications. *Journal of interactive Marketing* , 18 (3), 60-72.
27. Madhulatha, T. S. (2012). An overview on clustering methods. *Journal of Engineering* , 2 (4), 719-725.

- systems. *Journal of marketing management* , 18 (1-2), 193-219.
41. Wu, J., Shi, L., Lin, W.-P., Tsai, S.-B., Li, Y., Yang, L., et al. (2020). An Empirical Study on Customer Segmentation by Purchase Behaviors Using a RFM Model and K-Means Algorithm. *Mathematical Problems in Engineering* , 1-7.
42. Yeh, I.-C., Yang, K.-J., & Ting, T.-M. (2009). Knowledge discovery on RFM model using Bernoulli sequence. *Expert Systems with Applications* , 36 (3), 5866-5871.
43. Zeithaml, V. A., Rust, R. T., & Lemon, K. N. (2001). The customer pyramid: creating and serving profitable customers. *California management review* , 43 (4), 118-142.
44. Zhang, S., Zhang, C., & Yang, Q. (2003). Data preparation for data mining. *Applied artificial intelligence* , 17 (5-6), 375-381.
45. Zhang, Y., Bradlow, E. T., & Small, D. S. (2015). Predicting customer value using clumpiness: From RFM to RFMC. *Marketing Science* , 34 (2), 195-208.
- value. *WSEAS transactions on information Science and applications* , 6 (2), 319-328.
36. Shih, Y.-Y., & Liu, C.-Y. (2003). A method for customer lifetime value ranking—Combining the analytic hierarchy process and clustering analysis. *Journal of Database Marketing & Customer Strategy Management* , 11 (2), 159-172.
37. Singh, S. S., & Jain, D. C. (2017). Measuring customer lifetime value: models and analysis. *Review of Marketing Research* , 37-62.
38. Singh, S. S., Borle, S., & Jain, D. C. (2009). A generalized framework for estimating customer lifetime value when customer lifetimes are not observed. *Qme* , 7 (2), 182-205.
39. Smith, W. R. (1956). Product differentiation and market segmentation as alternative marketing strategies. *Journal of marketing* , 21 (1), 3-8.
40. Wilson, H., Daniel, E., & McDonald, M. (2002). Factors for success in customer relationship management (CRM)