منشورات جامعة وادي النيل
مجلة النيل للآداب والعلوم الانسانية
(ISSN: 1858 – 7054)
المجلد السادس، العدد الأول ،2025م
nilevalley.edu.sd

التنبؤ بالعوائد السكنية لدعم اتخاذ القرار باستخدام خوارزمية الغابة العشوائية: دراسة حالة محلية أمبدة

محمد ابراهيم محمد الحسن إبراهيم

كلية الدراسات العليا- جامعة النيلين

المؤلف: iflatoon44@gmail.com

المستخلص

تهدف هذه الدراسة إلى تحليل العوائد السكنية في منطقة محلية أمبدة باستخدام تقنيات تنقيب البيانات، مع التركيز على تطبيق خوارزميات شجرة القرار، لاستخلاص الأنماط والعوامل المؤثرة في قيمة العوائد. تم استخدام برنامج (WEKA) لتطبيق خوارزمية (Random Forest)، وتم تحليل أداء هذه الخوارزمية. أظهرت النتائج أن خوارزمية (Random Forest) حققت أفضل أداء من حيث الدقة والقدرة التنبؤية. توصلت الدراسة إلى أن العوامل المؤثرة في العوائد السكنية تشمل توفر الماء والكهرباء وبقية الخدمات. توصي الدراسة باستخدام نماذج شجرة القرار كأداة داعمة في عمليات اتخاذ القرار الاستثماري والتخطيط العمراني في المنطقة، لما توفره من قدرة على تفسير العلاقات بين المتغيرات والتنبؤ بالعوائد بدقة عالية.

كلمات مفتاحية: *شجرة القرار، (WEKA)، التنبؤ، تحليل البيانات، اتخاذ القرار.*

Predicting Residential Returns to Support Decision-Making Using the Random Forest Algorithm: A Case Study of Umbadda Locality

Abstract:

This study aims to analyze residential returns in Umbadda locality using data mining techniques, with a focusing on applying decision tree algorithms to extract patterns and identify the factors influencing return values .The (WEKA) software was used to apply the (Random Forest) algorithm, and its performance was analysed . The results showed that the (Random Forest) algorithm achieved the best performance in terms of accuracy and predictive capability. The study found that the key factors affecting residential returns include the availability of water, electricity, and other services .The study recommends the use of decision tree models as a supportive tool in investment decision-making and urban planning in the area, due to their ability to interpret relationships between variables and accurately predict returns.

Keywords*: Decision Tree, WEKA, Prediction, Data Analysis, Decision Making.*

مقدمة:

تعتبر العوائد السكنية من المؤشرات الرئيسية في تقييم الاستثمارات العقارية، حيث تعتمد على عوامل متعددة مثل الموقع، البنية التحتية، والعرض والطلب. في هذه الورقة يتم تطبيق نموذج شجرة القرار لتحليل العوائد السكنية في منطقة محلية أمبدة، بهدف تقديم رؤى دقيقة تساعد أصحاب القرار في تحسين قراراتهم. يعتمد التحليل على بيانات واقعية، مع استخدام تقنيات متقدمة لتنقيب البيانات لتحسين دقة التنبؤ.

مشكلة الدراسة: تتلخص المشكلة في عدم وجود آلية علمية موثوقة تطور وتحسن من اتخاذ القرار في منح تصاديق البناء، وايضا وضع الموازنات الايرادية الخاصة بالعوائد السكنية المستقبلية وامكانية التنبؤ بهذه العوائد.

اهداف الدراسة: تهدف هذه الدراسة إلى تطوير وتحسين عملية اتخاذ القرار في محلية أمبدة، من خلال التنبؤ بإيرادات العوائد السكنية، وتحديد العوامل المؤثرة فيها، وذلك باستخدام خوارزمية (Random Forest).

منهجية البحث: تم اتباع المنهج الوصفي التحليلي في مساحة مقدرة من هذه الدراس، والمنهج شبه التجريبي لما تقتضيه هذه الدراسة من إجراء التجربة العملية بمحلية أمبدة، حيث يتكون من المراحل التالية:

أولاً: تحديد المشكلة وأهداف البحث

- تحديد العوامل المؤثرة على رفع قيمة العوائد السكنية.
- استخدام تقنيات تنقيب البيانات لتطوير نموذج تنبؤي يساعد في اتخاذ القرار.

ثانياً: جمع البيانات

- الحصول على بيانات العوائد السكنية من الجهات الرسمية المختصة بمحلية أمبدة، حيث تحصلنا على عينة تتكون من 361 سجل منازل، يحتوي على بيانات رقمية وحرفية، وعينة اختبار لعدد 100 منزل اخرى لإجراء عملية التنبؤ بالعوائد السكنية.

ثالثاً: معالجة البيانات

- تنظيف البيانات من القيم المفقودة أو المتطرفة لضمان دقة التحليل، ومثال لذلك حذف كل القيم الشاذة والمتطرفة من عينة البيانات، باستخدام خاصية التعديل في البيانات المدمجة ببرنامج ويكا.

رابعاً: اختيار وتطبيق تقنيات تنقيب البيانات

- استخدام خوارزمية التنبؤ (Random Forest) وتم ضبط عدد الاشجار ليكون 15 شجرة للخوارزمية.
- تقييم أداء النماذج باستخدام مقاييس مثل متوسط الخطأ المطلق (MAE)، وتم استخدام مؤشرات مثل الخطأ المطلق المتوسط (MAE) والخطأ التربيعي المتوسط (RMSE) ومعامل التحديد (R^2) لتقييم أداء النماذج التنبؤية، حيث تساعد هذه المؤشرات في قياس دقة النتائج، ومقارنة أداء الخوارزميات، والتحقق من فعالية التحسينات المطبقة، كما تُعد أدوات معيارية تعزز من موثوقية النتائج ودعم اتخاذ القرار بشكل عملي.

خامسا: تحليل النتائج وتفسيرها

سادسا: تقديم التوصيات واتخاذ القرار

1. الدراسات السابقة:

هنالك العديد من الدراسات ذات صلة مباشرة أو غير مباشرة بموضوع الدراسة، والتي أمكن الحصول عليها اتفقت معنا في المنهجية والموضوع، واختلفت بعضها في ادوات التنفيذ، حيث تناولت الدراسات مواضيع الذكاء الاصطناعي وتنقيب البيانات ودورها في تطوير وتحسين اتخاذ القرار. الجدول التالي يوضح ذلك:

الدراسة/ المرجع	الهدف / ملخص	المنهج / النتائج / التوصيات	أوجه الشبه	أوجه الاختلاف
أرمان زارماجامبيتوف وآخرون(2019)	تقدم مقارنة بين خوارزمية (TAO) الجديدة وخوارزميات تقليدية من حيث الدقة وحجم الشجرة.	أظهرت (TAO) دقة أعلى غالبًا .يوصى باستخدام (TAO) للتطبيقات التي تتطلب دقة عالية.	الهدف المشترك هو تحسين الأداء.	الدراسة تقارن بين خوارزميات، بينما دراستنا تقوم باختبار خوارزمية ال (Random Forest) وتطبيقها على بيانات حقيقية عن العوائد السكنية.
ديانوي تشي وآخرون (2024)	تحسين خوارزمية (ID3)باستخدام قيم خبرة ميدانية وتبسيط الحسابات.	زادت كفاءة الخوارزمية دون التأثير على الدقة .يوصى بها في التصنيف السريع والفعال.	التركيز على تحسين خوارزمية شجرة القرار.	التحسين تم على (ID3)، بينما دراستنا تقوم باختبار خوارزمية ال (Random Forest) وتطبيقها على بيانات حقيقية عن العوائد السكنية.
تي .ميراندا لاكشمي وآخرون(2013)	تحليل أداء ID3) و C4.5و (CART باستخدام بيانات طلابية.	(CART)أظهرت دقة أعلى .يوصى باختيار الخوارزمية حسب طبيعة البيانات.	الاهتمام بتحسين الأداء.	تركز على خوارزميات تقليدية وبيانات تعليمية، بينما دراستنا تقوم باختبار خوارزمية ال (Random Forest) وتطبيقها على بيانات حقيقية عن العوائد السكنية.

جدول رقم 1: مقارنة بين الدراسة والدراسات السابقة لها

2. **الخلفية النظرية والادوات المستخدمة:**

أ. **التطوير الإداري (بيتر، 2021)**

التطوير الإداري هو عملية مستمرة، تهدف إلى تحسين الأداء التنظيمي، وزيادة كفاءة المؤسسات وتعتمد فعاليته على التزام القيادة تبني التكنولوجيا، وتطوير الموارد البشرية، ومع استمرار التحديات، فإن الابتكار والتخطيط الاستراتيجي يمثلان مفتاح النجاح في تنفيذ التطوير الإداري بكفاءة.

ب. **مفهوم اتخاذ القرار وأهميته** (يوسف، ي، 2019)

اتخاذ القرار هو عملية اختيار البديل الأفضل من بين عدة بدائل متاحة، بناءً على تحليل البيانات والتنبؤ بالنتائج، وتقييم المخاطر.

يعتبر اتخاذ القرار عنصرًا أساسيًا في نجاح أي مؤسسة، حيث يؤثر بشكل مباشر على كفاءة الأداء، وجودة الخدمات، وتحقيق الأهداف الاستراتيجية.

ت. **تنقيب البيانات (Data Mining)** (السيد، أ. 2020):

أولاً: تعريف تنقيب البيانات

تنقيب البيانات هو العملية التي تهدف إلى اكتشاف الأنماط والمعرفة، من خلال تحليل كميات ضخمة من البيانات غير المنظمة أو المنظمة، باستخدام تقنيات الذكاء الاصطناعي، والتعلم الآلي والإحصاء.

يتضمن تنقيب البيانات مجموعة من الخطوات مثل:

1. التصفية: تنقية البيانات وتجهيزها للتحليل.

2. استكشاف البيانات: تحليل البيانات لاستخلاص الأنماط.

3. التفسير: تقديم التفسيرات للأنماط المستخلصة.

ثانياً: أهمية تنقيب البيانات

في عصر البيانات الضخمة، أصبحت الشركات والمؤسسات تعتمد بشكل كبير على تنقيب البيانات من أجل اتخاذ قرارات أكثر دقة، استخدام هذه التقنية يساعد في:

- تحقيق فعالية الأعمال: من خلال تحسين استراتيجيات التسويق، والبيع، والإنتاج.
- التنبؤ: يساعد في التنبؤ بالاتجاهات المستقبلية مثل سلوك العملاء.
- الكشف عن الأنماط المخفية: مثل العمليات الاحتيالية أو الأنماط المرضية في بيانات الرعاية الصحية.

ثالثاً: دورتنقيب البيانات في تطوير اتخاذ القرار

تنقيب البيانات هو عملية تحليل كميات ضخمة من البيانات لاكتشاف الأنماط، العلاقات، والتوجهات المخفية التي قد تساهم في اتخاذ قرارات مدروسة، و يعمل تنقيب البيانات على تحسين عملية اتخاذ القرار بعدة طرق:

أ. اكتشاف الأنماط المخفية: تنقيب البيانات يساعد في اكتشاف الأنماط المخفية، التي قد تكون غير واضحة، أو يصعب ملاحظتها باستخدام الطرق التقليدية من خلال تحليل البيانات التاريخية، ويمكن اكتشاف العلاقات التي قد تساهم في تحسين القرارات المستقبلية.

مثال: في التجارة، من خلال تحليل سلوك العملاء، يمكن اكتشاف أنماط معينة تؤدي إلى زيارات أو مبيعات أكثر، مما يساعد في تحديد استراتيجيات التسويق الأفضل.

ب. اتخاذ قرارات بناءً على البيانات: وذلك باستخدام تقنيات مثل التصنيف والتجميع، يمكن لعملية تنقيب البيانات تقسيم البيانات إلى فئات أو مجموعات، مما يساعد في اتخاذ قرارات قائمة على فهم دقيق للبيانات.

مثال: في القطاع المالي يمكن لتقنيات تنقيب البيانات تحليل المعاملات المالية، والكشف عن الأنشطة الاحتيالية، مما يساعد المؤسسات المالية في اتخاذ قرارات سريعة للحد من المخاطر.

ج. التنبؤ بالاتجاهات المستقبلية: من خلال نمزجه البيانات واستخدام الخوارزميات التنبؤية يمكن لتنقيب البيانات أن يساهم في التنبؤ بالاتجاهات المستقبلية، وتحليل كيف يمكن أن تؤثر المتغيرات على قرارات معينة.

مثال: في مجالات مثل إدارة المخزون، تساعد تقنيات تنقيب البيانات الشركات على التنبؤ بالطلب المستقبلي على المنتجات، مما يسهل اتخاذ قرارات حول الإنتاج والمخزون.

د. تحسين استراتيجيات الأعمال: من خلال تحليل البيانات الضخمة باستخدام تقنيات التنقيب، يمكن للشركات تحديد الاستراتيجيات الأكثر فعالية، واتخاذ قرارات مدروسة بشأن تخصيص الموارد، وتحديد الأسعار، وتحسين العمليات.

مثال: يمكن لتقنيات تنقيب البيانات في شركات البيع بالتجزئة تحليل سلوك العملاء وتفضيلاتهم، مما يساعد في تحسين استراتيجيات العروض الترويجية وتخصيص المنتجات.

ث. برنامج ويكا للتنقيب في البيانات (فتيحه بوهرين،2015)

تعتبر (WEKA) وأحدة من أكثر الأدوات شيوعًا في مجال تعلم الآلة وتنقيب البيانات بفضل سهولتها وفعاليتها. تتيح هذه الأداة للمستخدمين تجربة خوارزميات مختلفة، وتقديم تقارير مفصلة عن الأداء، وفحص البيانات وتجهيزها للمعالجة.

ج. المفهوم العام والاساس الرياضي لخوارزمية (Random Forest) (.(2012) .Han, J., Kamber, M., & Pei, J)

أولاً: مقدمة عن (Random Forest)

خوارزمية (Random Forest) هي تقنية تعلم آلي تعتمد على تجميع عدة أشجار قرار لتحسين أداء التصنيف أو التنبؤ. تُعتبر من أساليب التعلم الجماعي (Ensemble Learning) التي تهدف إلى تقليل التحيز والاختلاف في النموذج النهائي مقارنة باستخدام شجرة قرار واحدة فقط.

ثانياً: توليد الأشجار

يتم توليد كل شجرة قرار في الغابة باستخدام طريقتين رئيسيتين لتعزيز التنوع:

- اختيار عينة بيانات (Bootstrap) أي: سحب عينة عشوائية مع الإرجاع من مجموعة التدريب.
- اختيار مجموعة عشوائية من السمات عند كل عقدة لبناء الشجرة.

ثالثاً: الإخراج والتجميع

إذا افترضنا أن لدينا مجموعة بيانات تدريبية تتكون من N عينة

$D = \{(x_1, y_1), (x_2, y_2), \ldots,$

$(x_N, y_N)\},$

حيث $x_i \in \mathbb{R}^d$ ، و y_i هو ناتج التصنيف أو الانحدار.

يتم تدريب عدد M من أشجار القرار $\{T_1(x), T_2(x), \ldots, T_M(x)\}$ ، بحيث كل شجرة T_m يتم تدريبها باستخدام عينة (Bootstrap) وعشوائية السمات.

بالنسبة للتصنيف، يكون الناتج النهائي:

$$RF(x) = \text{mode}\{T_1(x), T_2(x), \ldots, T_M(x)\}$$

وبالنسبة للانحدار، يكون الناتج:

$$RF(x) = (1/M) * \sum T_m(x) \text{ حيث } m = 1 \text{ إلى } M$$

رابعا تقدير الخطأ وتقنية ((Out-Of-Bag (OOB)

يُستخدم خطأ العينات غير المستخدمة في التدريب (Out-Of-Bag Error) لتقدير دقة النموذج دون الحاجة لمجموعة اختبار مستقلة . كل شجرة يتم تدريبها باستخدام عينة (Bootstrap)، وبالتالي فإن حوالي 1/3 من العينات لا تُستخدم في التدريب .يمكن استخدام هذه العينات لتقييم أداء النموذج.

خطأ (OOB)يُحسب كالتالي:

$$\text{OOB Error} = (1/N) * \sum I(y_i \neq \hat{y}_i^{\{OOB\}})$$

حيث I هي دالة المؤشر.

خامساً: المزايا الرياضية

- تقليل التباين (Variance Reduction) من خلال التجميع.
- الحفاظ على انخفاض التحيز (Low Bias) بالمقارنة مع النماذج الخطية.
- تقليل الإفراط في التخصيص (Overfitting) مقارنة بشجرة القرار الفردية .

ح. الفجوة البحثية

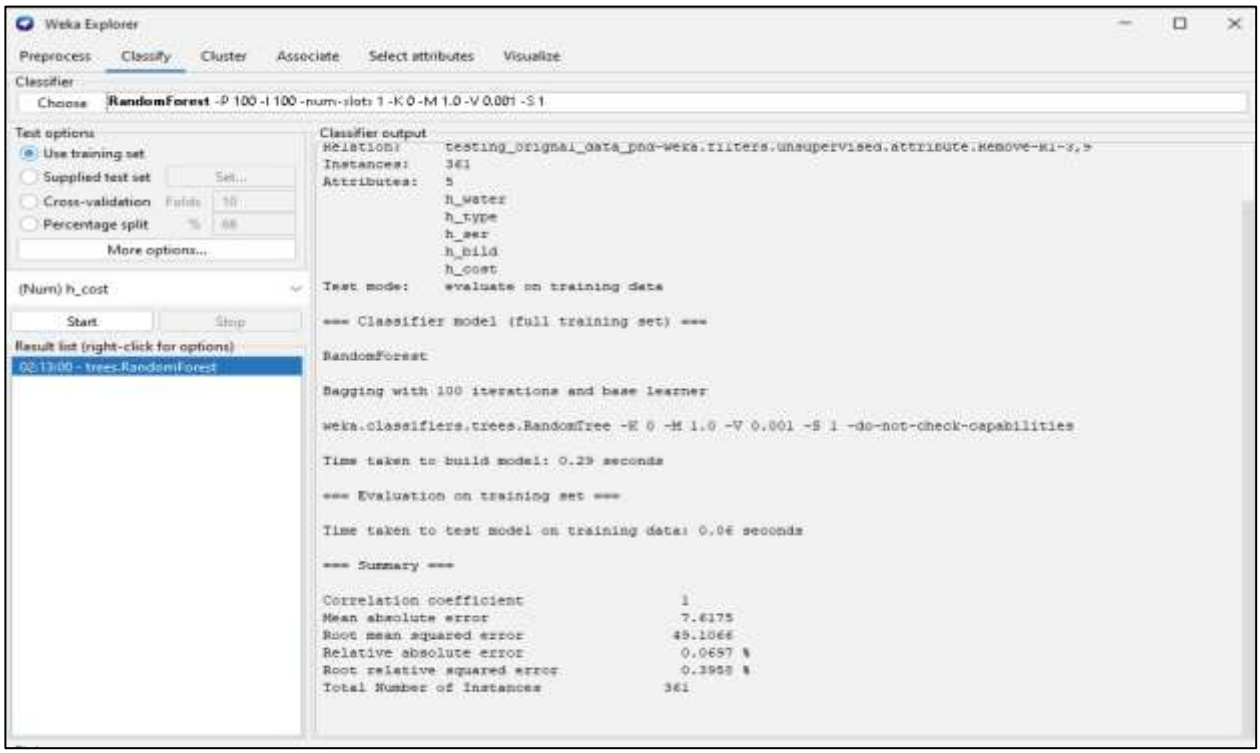
هنالك العديد من الفجوات البحثية في هذا المجال، تهدف هذه الدراسة الي سدها، مثل:

- العوامل التي تؤثر علي زيادة العوائد السكنية بالمحلية.
- الطريقة العلمية في منح تصاديق البناء ونوعها بالمحلية.
- الطريقة العلمية الموثوقة لتوزيع الخدمات بالمحلية.

3. سبب استخدام خوارزمية (Random Forest) في هذه الدراسة

سبب استخدام خوارزمية ال Random Forest)) في هذه الدراسة لأنها تتميز بقدرتها على التعامل مع البيانات الكبيرة والمتعددة الأبعاد، كما تُعتبر فعالة في تقليل التحيز وزيادة التباين، مما يجعلها مناسبة للتطبيقات التنبؤية المختلفة.

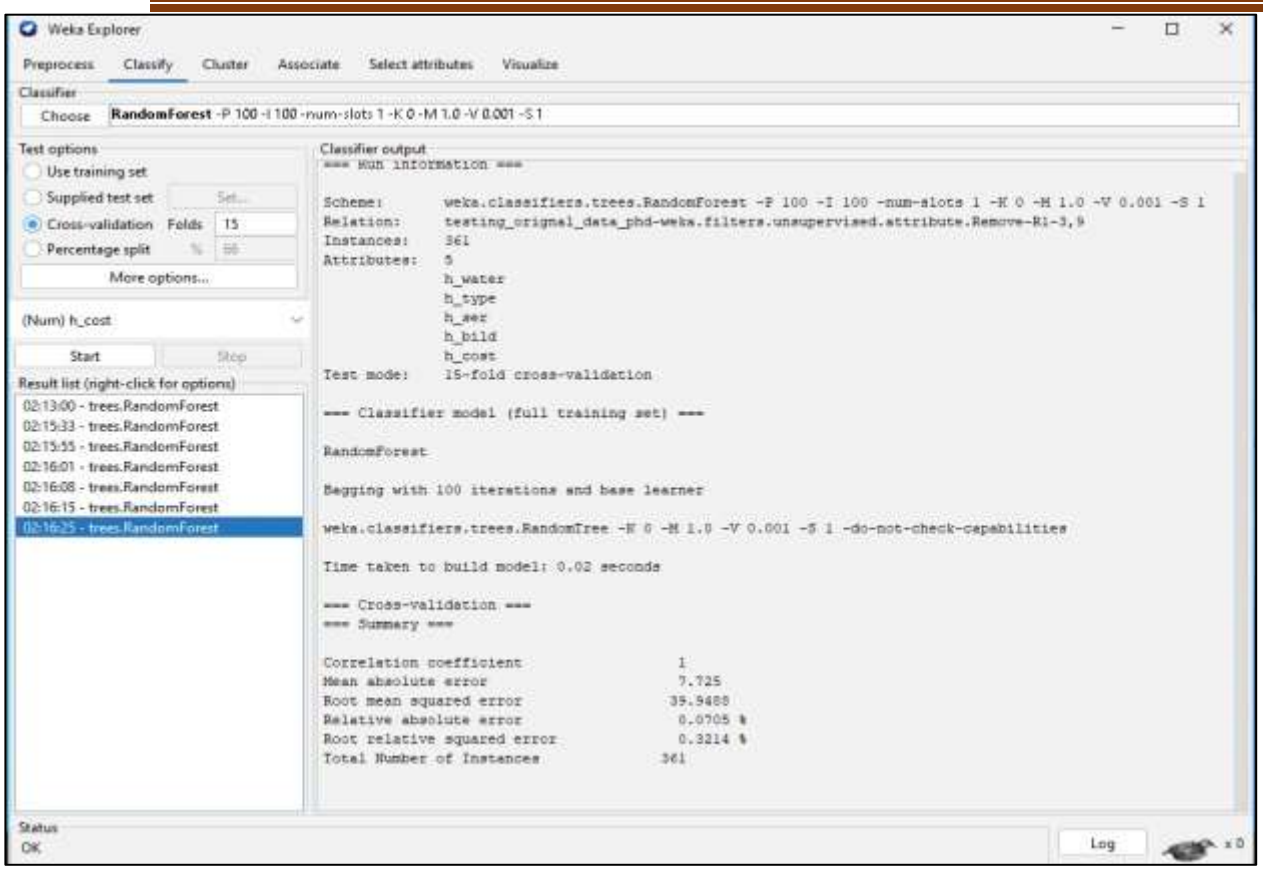
4. تدريب وتقييم الخوارزمية على عينة البيانات التاريخية بالمحلية ومناقشة النتائج



شكل رقم 1: شاشة عرض ملخص عمل الخوارزمية

مناقشة وتحليل النتائج في شكل 1

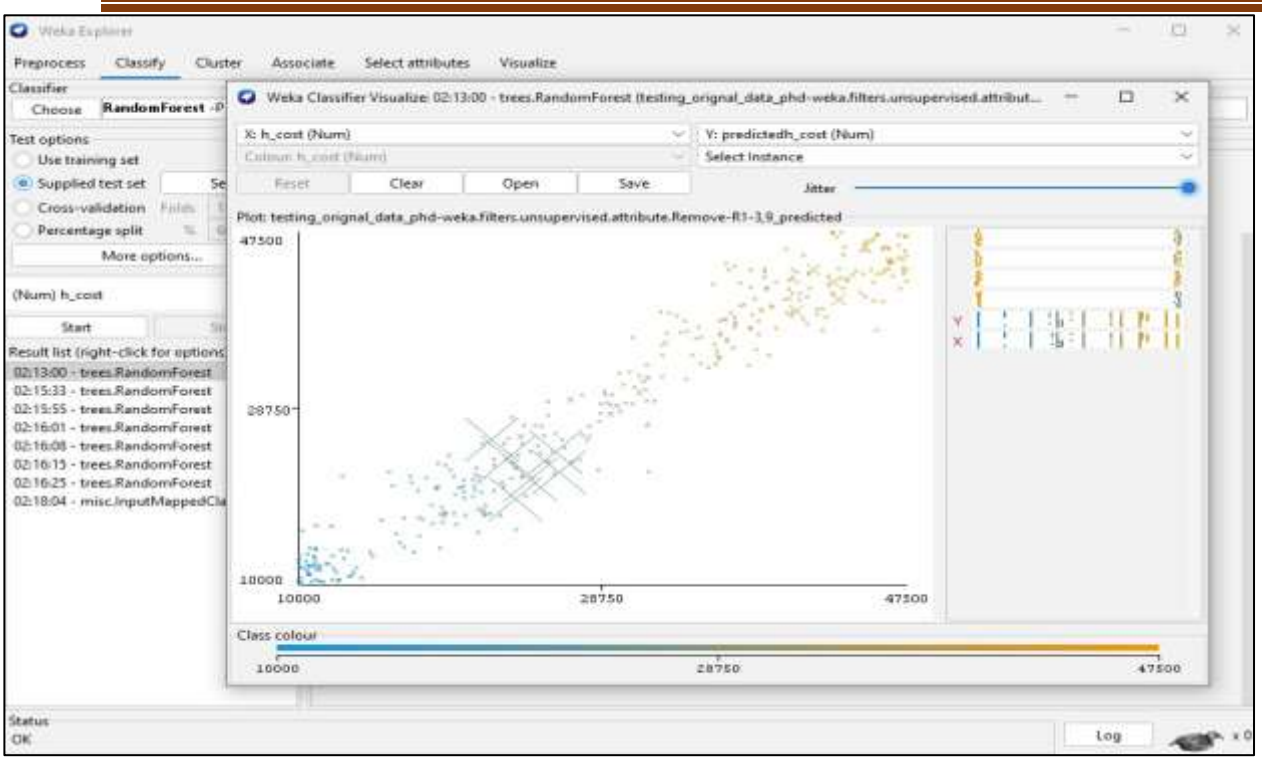
- بعد تشغيل الخوارزمية على البيانات الاصلية ظهر لنا في النتائج ان معامل الارتباط يساوي الواحد الصحيح، وهو المعامل الذي يقيس مدي قوة العلاقة بين القيم الحقيقية والقيم المتوقعة، حيث تتراوح قيمته عادة بين -1 و +1، وكلما اقتربت قيمته من الواحد كان النموذج دقيق جداً، وإذا كانت مساوية للصفر فذلك يعني أن النموذج عشوائي ولا يتنبأ بشكل جيد وإذا كانت قريبه من -1 فإن النموذج يعكس العلاقة بشكل خاطئ وعليه فإن النموذج هنا لديه دقة جيدة جداً في التنبؤ بالقيم العددية.
- وإذا نظرنا إلى الخطأ المطلق المتوسط (Mean Absolute Error - MAE) والذي هو عبارة عن متوسط الفرق المطلق بين القيم الحقيقية والقيم التي توقعها النموذج الذي بدوره كلما كان صغيراً كان النموذج أفضل، وجدناه هنا يساوي (7.7615) وذلك يعني أن متوسط خطأ التوقعات هو (7.7615) عن القيم الحقيقية وهذه تعتبر نتيجة جيدة.
- وإذا نظرنا إلى الجذر التربيعي لمتوسط مربعات الخطأ (Root Mean Squared Error - RMSE) والذي بدوره يقيس مدى تباعد القيم المتوقعة عن القيم الحقيقية، مع اعطاء أهمية أكبر للأخطاء الكبيرة، وعندما يكون ال (RMSE) مرتفعاً فقد يكون هنالك خطأ كبير في بعض التوقعات الفردية، وهنا نجده ايضا مساويا (49.1066) مما يدل على أن انحراف القيم المتوقعة عن الحقيقية هو (49.1066) وتعتبر نتيجة جيدة لأن الخطأ لا يكون كبيراً إن وجد.
- وإذا نظرنا إلى نسبة الخطأ المطلق النسبي (Relative Absolute Error - RAE) والتي تقارن بين ال (MAE) وخطأ نموذج بسيط يعتمد فقط على المتوسط العام للبيانات، وإذا كان ال (RAE) قريبا من صفر% فالنموذج أفضل من التوقعات العشوائية، وعندنا هنا يعتبر أفضل لأنه جاء 0697% وهو قريب جداً من الصفر.
- وإذا نظرنا الي نسبة الجذر التربيعي لمتوسط مربعات الخطأ النسبي (Root Relative Squared Error - RRSE)، وهو مشابه ايضا لل (RMSE) لكنه يحسب كنسخة نسبية مقارنة بنموذج بسيط يعتمد على المتوسط، وكلما قلّ كان النموذج أفضل، وهنا جاء 0.3958 مما يدل على أن النموذج دقيق لحد كبير.
- اخيراً تظهر عدد العينات المستخدمة (Total Number Of Instances)، والتي كلما كانت كبيرة زادت دقة النموذج، بشرط أن تكون البيانات متنوعة وكافية.



شكل رقم 2: شاشة عرض ملخص عمل تقييم الخوارزمية

مناقشة وتحليل النتائج في شكل 2

- في هذه الشاشة تم استخدام أحد خيارات تقييم الخوارزمية وهو (Cross-Validation) الذي يقوم بتقسيم البيانات إلى تدريب واختبار، وعند ملاحظة القراءات في قسم الملخص ومقارنته مع ملخص الخوارزمية في الشكل 1 فأننا نلاحظ إن الاختلاف بسيط، ولا يؤثر كثيرا على عمل الخوارزمية، وهذا يدل إن عملية التقييم اثبتت فعالية الخوارزمية، وأنها صالحة وجاهزة لأي عملية استنتاج أو تنبؤ.



شكل رقم3: شاشة معاينة الاخطاء

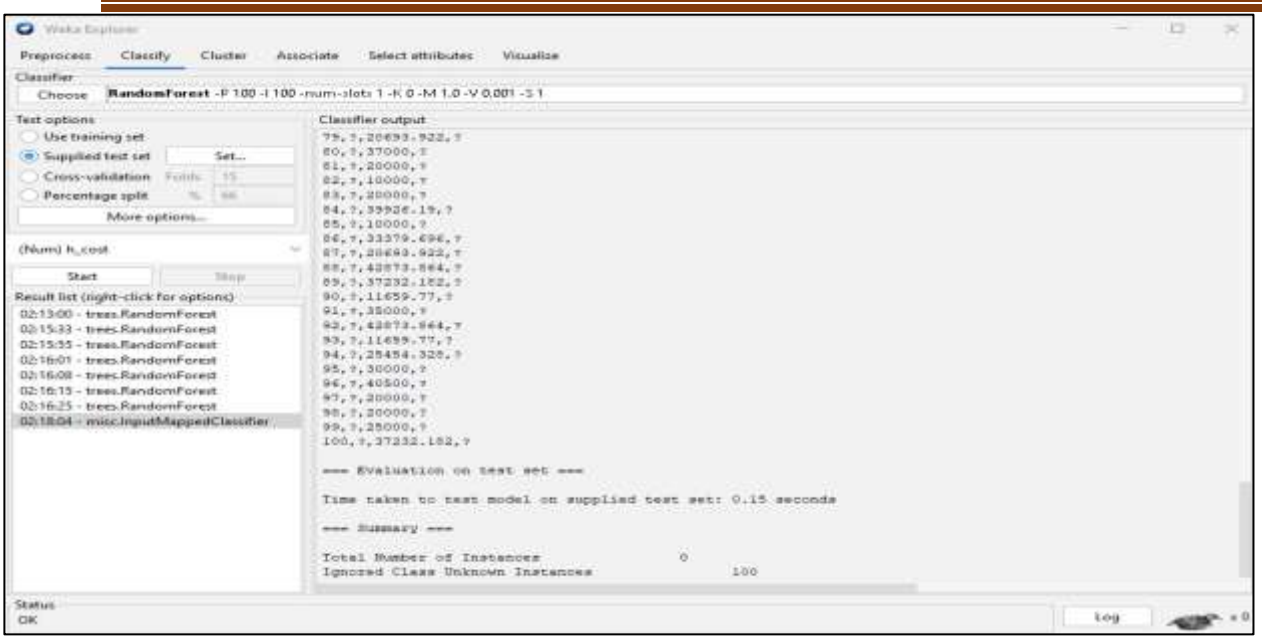
مناقشة وتحليل النتائج في شكل 3

- يتم عبرها معاينة عملية الاخطاء بعد تطبيق الخوارزمية، وملاحظة القيم الشاذة، وأيضاً عبرها يتم حفظ بيانات التنبؤ الجديدة في ملف خارجي.

وإذا نظرنا إلى الشكل جيداً فإننا نجد أنه لا توجد أي بيانات متطرفة على الرسم أو أي قيم شاذه. وكل القيم تتراوح بين فئات العوائد السكنية، وذلك إن دل فإنما يدل على شيئين:

الأول: أن البيانات تم تجهيزها جيداً من حيث القيم المتطرفة أو المفقودة.

الثاني: أن الخوارزمية قد أدت دورها بكل دقة في عملية التحليل.

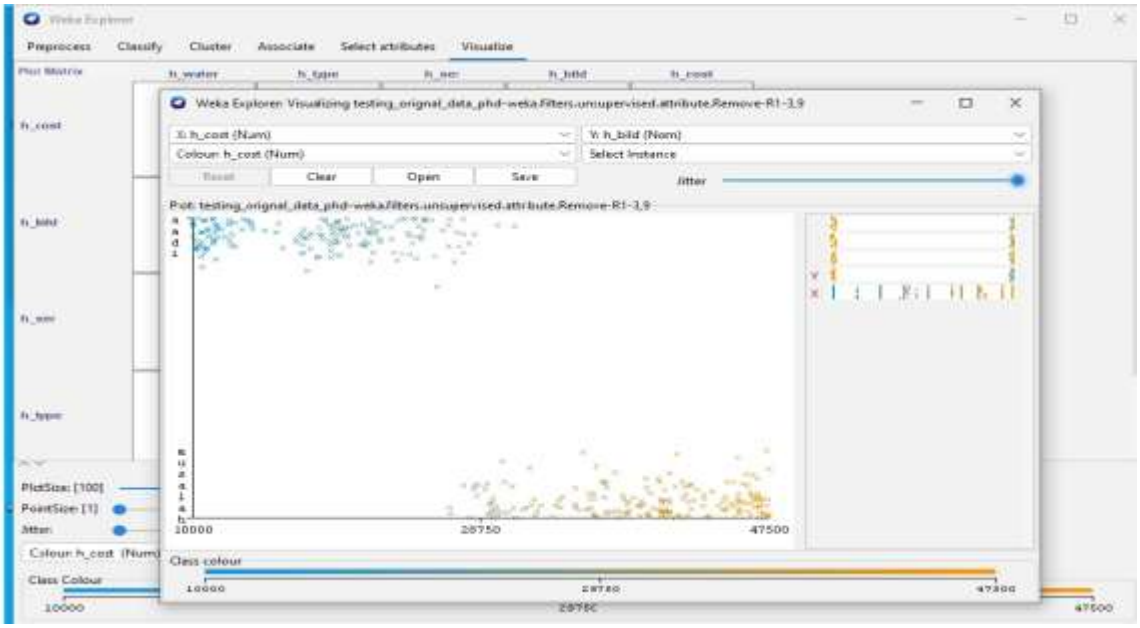


شكل رقم 4: شاشة اجراء التنبؤ

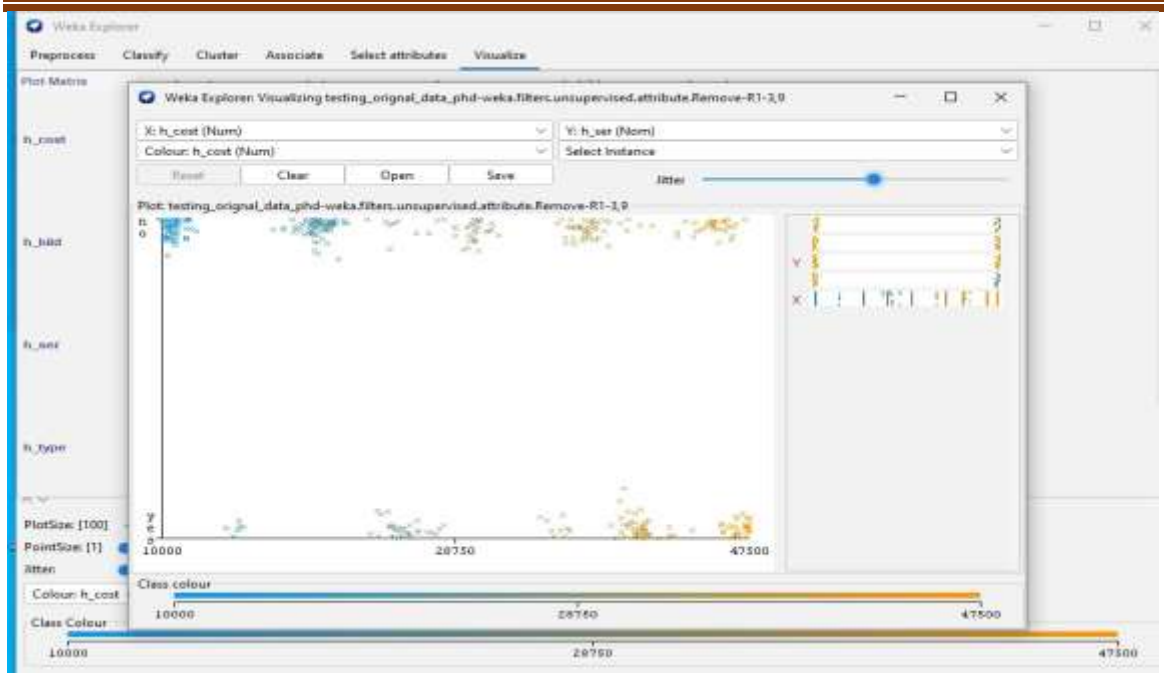
شكل رقم 5: ملف مخرجات التنبؤ.

مناقشة وتحليل النتائج في شكل 4 وشكل رقم 5

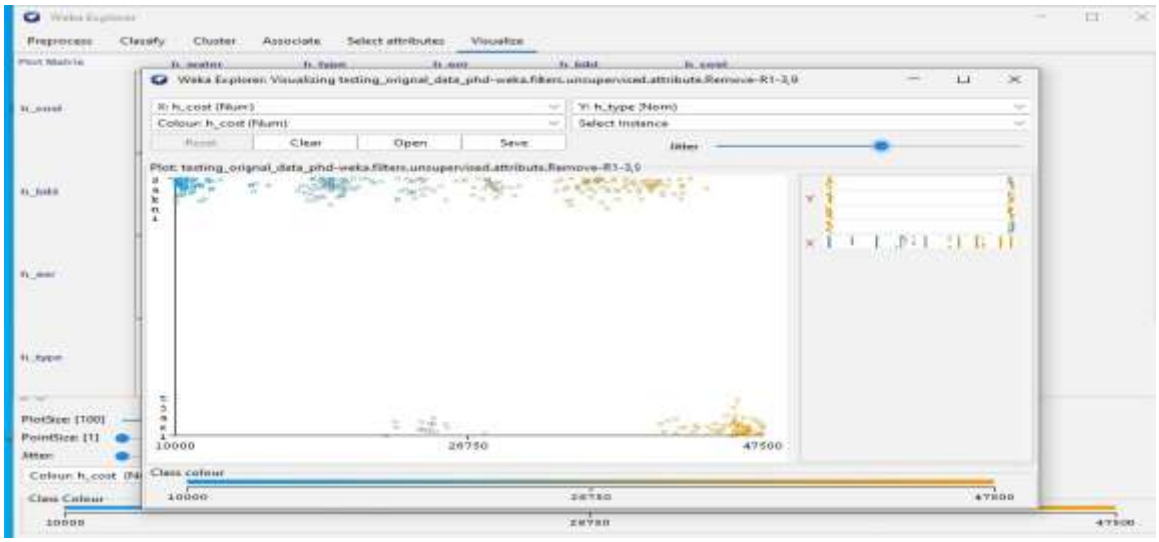
- بعد تنفيذ الخوارزمية واختبارها، قمنا بتشغيلها على عينة اختبار خارجية، تحتوي على مائة منزل غير محدد العوائد السكنية لهم، على اساس أن تقوم الخوارزمية بعملية التنبؤ بالقيم المطلوبة، وقد قامت باستنتاج كل القيم الجديدة داخل البرنامج في الشكل 4 لنقوم بحفظها بملف خارج البرنامج كما بالشكل 5، وعند تحليل هذه القيم الجديدة ومقارنتها مع القيم الحقيقية السابقة وجدنا أن هذه الاستنتاجات ممتازة مع وجود انحراف بسيط في بعض القيم.
- عند تحليل النتائج بالشكلين أعلاه، وجدنا أن هنالك ارتباطاً كبيراً ومؤثراً في قيمة العوائد السكنية، تقوم به بيانات توفر الخدمات مثل الماء والكهرباء، والمستشفيات وبقية المرافق، حيث أنه كلما كان المنزل قريب منهم فإن عوائده السكنية تكون عالية، وهذا تناسب طردي.



شكل رقم 6: مخطط العلاقة بين متغير نوع البناء وقيمة العوائد



شكل رقم 7: مخطط العلاقة بين متغير توفر الخدمات وقيمة العوائد



شكل رقم 8: مخطط العلاقة بين متغير الغرض من المنزل وقيمة العوائد

مناقشة وتحليل النتائج في شكل 6 وشكل رقم 7 وشكل رقم 8

- تلاحظ أن هنالك علاقة طردية بين متغير نوع البناء والغرض من البناء، وتوفر الخدمات مع قيمة العوائد السكنية، فكلما كان نوع البناء مسلح زادت قيمة العوائد، وأيضاً كلما كانت الخدمات متوفرة زادت قيمة العوائد، وأيضاً كلما كان الغرض تجاري زادت قيمة العوائد السكنية. عليه فإن هذه المتغيرات مرتبطة وتؤثر على قيمة العوائد بشكل مباشر.

خاتمة:

- تم استعراض كيفية تحسين عملية اتخاذ القرار باستخدام أدوات تنقيب البيانات ضمن إطار مشروع الحكومة الإلكترونية، مع التركيز على دراسة تطبيقية للعوائد السكنية في الوحدات الإدارية بمحلية أمبدة. أظهرت النتائج أن توظيف تقنيات تنقيب البيانات يسهم بشكل فعال في تحسين دقة القرارات، من خلال تحليل البيانات الضخمة واستنباط الأنماط والعلاقات التي قد لا تكون واضحة بالأساليب التقليدية. كما أكدت الدراسة أهمية تبني هذه الأدوات في المؤسسات الحكومية لتعزيز كفاءة السياسات العامة، لا سيما في مجالات التخطيط العمراني والإدارة المحلية. بناءً على النتائج المتوصل إليها، توصي الدراسة بتوسيع نطاق استخدام تقنيات تنقيب البيانات في مختلف القطاعات الحكومية، مع توفير البنية التحتية اللازمة ودعم الكوادر البشرية بالتدريب المتخصص. كما يمكن أن تمثل هذه الدراسة نقطة انطلاق لأبحاث مستقبلية تستكشف مجالات أخرى من الخدمات الحكومية الإلكترونية، مما يسهم في تعزيز فاعلية نظم المعلومات الإدارية ودعم تحقيق التنمية المستدامة.

آفاق البحث المستقبلي:

- تجربة مقارنة أداء ((Random Forestمع خوارزمية أخرى مثل (XGBoost أو .(Gradient Boosting
- تطوير لوحة معلومات ((Dashboardتعرض نتائج النموذج لصانع القرار.
- إضافة متغيرات مكانية باستخدام نظم المعلومات الجغرافية (.(GIS

النتائج:

- تأثير توفر الخدمات على العوائد السكنية: أظهرت النتائج أن زيادة الإيرادات من العوائد السكنية تتناسب طرديًا مع توفر الخدمات العامة، مثل المستشفيات، مراكز الشرطة، المكاتب المدنية، إضافة إلى خدمات المياه والكهرباء، وعليه فإن تحسين وتوسيع البنية التحتية في المناطق السكنية يؤدي إلى زيادة العوائد السكنية بشكل ملحوظ.
- ضوابط منح تصاريح البناء: تشير النتائج إلى أن منح تصاريح البناء في المناطق التي تفتقر إلى الخدمات الأساسية، مثل المياه والكهرباء يؤدي إلى انخفاض العوائد السكنية بشكل كبير، مما يجعل تحسينها مستقبلاً أكثر صعوبة. لذلك توصي الدراسة بعدم منح تصاريح البناء في مثل هذه المناطق إلا بعد ضمان توفر الخدمات الأساسية فيها.
- أولوية التنمية للمناطق غير المخدومة: تؤكد النتائج أن توجيه التنمية نحو المناطق التي تفتقر إلى الخدمات الأساسية، من خلال إدخال المياه والكهرباء إليها، يسهم في تنشيط هذه المناطق وزيادة الطلب على تصاريح البناء، مما يؤدي بدوره إلى ارتفاع الإيرادات المتحصلة من العوائد السكنية.

التوصيات

- مستقبلية: دمج بيانات زمنية لتقييم تغير العوائد السكنية بمرور الوقت.
- منهجية: تجربة نماذج هجينة تجمع بين شجرة القرار والشبكات العصبية.
- تطبيقية: تطبيق النموذج على مناطق أخرى للتحقق من قابلية التعميم.

المراجع

السيد، أ. (2020). 'تحليل وتنقيب البيانات في نظم المعلومات'. الطبعة الثانية. القاهرة: دار الفكر.

فتيحة بوهرين، ف، (2015)، 'ورقة عمل عن التنقيب في البيانات باستخدام ويكا'، جامعة قسنطينة- الجزائر.

بكة، م. (2023) 'تنقيب البيانات: المراحل والأدوات والخوارزميات والكتب'.

بيتر دراكر، ب، (2021)، 'التطوير الإداري الحديث'.

يوسف، ي، (2018). 'مقدمة في نظم المعلومات الإدارية'. الرياض: دار الخريجي.

Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann.