

التنقيب في البيانات و اتخاذ القرارات (نموذج تطبيقي لخزان خشم القرية)

Data mining and decision-making (An Applied Study on KhashmAlgirba Dam)

د. سيف الدين عثمان فتوح (أستاذ علوم الحاسوب المشارك)

(بكلية الإمارات للعلوم والتكنولوجيا)

د. الشفييع جعفر محمود (أستاذ مساعد)

(جامعة المجمعة، كلية العلوم الإنسانية بالغاظ، قسم تقنية المعلومات)

المستخلص:

إن استخدام تقنيات التنقيب في البيانات يوفر للمؤسسات في جميع المجالات القدرة على استكشاف و التركيز على أهم المعلومات في قواعد البيانات، كما تركز تقنيات التنقيب في البيانات كذلك على بناء التنبؤات المستقبلية و استكشاف السلوك و الاتجاهات مما يسمح باتخاذ القرارات الصحيحة و في الوقت المناسب.

كلمات مفتاحيه: تنقيب البيانات ، التنبؤ ، البيانات

Abstract

The use of techniques Data mining provides for institutions in all areas the ability to explore and focus on the most important information in the databases, also focusing techniques data mining as well as to build a future forecasts and exploration behavior and trends, which allows to take the right decisions and taken in a timely manner.

Keywords: Data mining, Forecasting, Data

مقدمة:

يتميز عصرنا الراهن (عصر الانترنت و الاقتصاد الرقمي) بالسييل العظيم و الانتشار الواسع النطاق للبيانات حتى أصبح من المستحيل على المحللين استخلاص معلومات ذات معنى باللجوء فقط إلى المداخل التقليدية للتحليل التمهيدي للبيانات.

مع وجود كميات كبيرة من البيانات المخزنة في قواعد البيانات و مخازن البيانات ازدادت الحاجة إلى تطوير أدوات تمتاز بالقوة لتحليل البيانات و استخراج المعلومات و المعارف منها، من هنا ظهر ما يسمى بالتنقيب في البيانات كتقنية تهدف إلى استخراج المعرفة من كميات هائلة من البيانات. و هي تقنية حديثة فرضت نفسها بقوة في عصر المعلوماتية، و استخدامها يوفر للشركات و المنظمات في جميع المجالات القدرة على استكشاف و التركيز على أهم المعلومات في قواعد البيانات، كما تركز تقنيات التنقيب على بناء التنبؤات المستقبلية و استكشاف السلوك و الاتجاهات مما يسمح باتخاذ القرارات الصحيحة و اتخاذها في الوقت المناسب. و التي تعتبر بدورها مرحلة من مراحل عملية أكثر تعقيدا هي استكشاف المعرفة في قواعد البيانات، و المرتبطة إلى حد بعيد بعملية تطوير أخرى مهمة جدا هي مستودعات البيانات. حيث أن الكثير من الشركات و المنظمات الرائدة اليوم تستخدم عملية استكشاف المعرفة في قواعد البيانات بشكل منهجي و منظم بوصفها تشكل جوهر العمل الذي يعتمد عليه في تفعيل النشاط و تحقيق الميزة التنافسية.

أهمية أسلوب تنقيب البيانات:

1) عملية تحليلية للاستكشاف والبحث في بيانات ضخمة وهائلة لاستخراج أنماط مفيدة وإيجاد العلاقات ومدى الارتباط بين عناصرها.

2) من اجل تحليل البيانات للحصول علي علاقات جديدة وغير متوقعة.

- 3) نقل عالم الأنظمة الواقعي إلى عالم افتراضي يمارس فيه متخذوا القرار التحليل واختبار الفرضيات علي شاشات الحاسوب ذات القدرة الرسومي العالية الدقة إلى أن يصلوا إلى ما يطمحون إليه من فهم وقناعات قبل اتخاذ القرارات بشأن ما يدرسونه من أنظمة.
- 4) التنبؤ ومن ثم استنتاج إجابات مقدرة تقديرا إحصائيا لكميات.

أهداف أسلوب تنقيب البيانات:

- 1) إن التنقيب في قواعد البيانات يهدف إلى انتزاع واستخلاص أنماط مفيدة، وهي تكنولوجيا حديثة، أصبحت مهمة في ظل التطور السريع وانتشار استخدام قواعد البيانات.
- 2) استخدامها يوفر للمؤسسات وأجهزة الأمن في جميع المجالات القدرة علي استكشاف، والتركيز علي أهم المعلومات في قواعد البيانات.
- 3) تركز تقنيات التنقيب علي بناء التنبؤات المستقبلية واستكشاف السلوك والاتجاهات، مما يسمح بتقدير القرارات الصحيحة واتخاذها في الوقت المناسب.
- 4) تجيب تقنيات التنقيب علي العديد من الأسئلة، وفي وقت قياسي، بخاصة تلك النوعية من الأسئلة التي يصعب الإجابة عليها، إن لم يكن مستحيلا، باستخدام تقنيات الإحصاء الكلاسيكية، والتي كانت إن وجدت فإنها تستغرق وقتا طويلا والعديد من الإجراءات.

الإطار النظري

أولاً: مفهوم التنقيب في البيانات (Data Mining)

ظهر مصطلح التنقيب في البيانات في منتصف التسعينات في الولايات المتحدة الأمريكية، وهو يجمع ما بين الإحصاء و تكنولوجيا ت الإعلام (قواعد البيانات، الذكاء الاصطناعي، التعلم الآلي « *Machine Learning* ... »).

و توجد عدة تعريفات لهذا المفهوم منها ، حيث يمكن تعريفها بأنها : " الاستكشاف الآلي أو المؤتمت لأنماط شائعة و غير جلية مخفية في قاعدة بيانات معينة " ⁽¹⁾ ، أو أنها: " إجراءات تحليل دقيقة وذكية،

تفاعلية و تسلسلية، تسمح لمسيرى النشاطات عند استخدام هذه الإجراءات باتخاذ قرارات والقيام بأعمال ملائمة في صالح النشاط المسؤولين عنه و المؤسسة التي يعملون بها" (2)، أو أنها: " عبارة عن تحليلات لكمية كبيرة من البيانات بغرض إيجاد قواعد و أمثلة و نماذج التي يمكن أن تستخدم تقود و تدل أصحاب القرار، و تتنبأ بالسلوك المستقبلي" (3)، كما يمكن تعريفها كذلك بأنها: " تحليل لمجموعات كبيرة الحجم من البيانات المشاهدة للبحث عن علاقات محتملة و تلخيص للبيانات في أشكال جديدة لتكون مفهومة و مفيدة لمستخدمها" (4).

من خلال التعريفات السابقة يمكن القول إن التنقيب في قواعد البيانات يهدف إلى استخراج المعلومات المخبأة فيها، واستخدامها يوفر للمؤسسات في جميع المجالات على استكشاف، والتركيز على، أهم المعلومات في قواعد البيانات، بالإضافة إلى كثرة البيانات الموجودة والمخزنة في ما يسمى بقواعد البيانات (*database*)، أصبحت موضوع تساؤل من عديد من الباحثين للاستفادة منها، ومع زيادة انتشار مستودعات التخزين الضخمة ما يدعى (*data warehouses*)، أصبح من الضروري إيجاد تقنيات وطرق ووسائل لاستخلاص المعلومات والمعرفة من مثل هذه البيانات المكثفة واستغلالها في حل المشاكل واتخاذ القرارات، باستخدام تطبيقات الحاسوب الحديثة والتي تعتبر تكنولوجيا حديثة ذكية قائمة على جعل الحاسوب "يفكر كما يفكر الإنسان ويفعل كما يفعل الإنسان" جاءت فكرة الكشف والتنقيب على هذه البيانات بطرق ذكية للمساعدة في حل المشاكل واتخاذ القرارات. وتعتبر خطوة من خطوات استكشاف المعرفة من قواعد البيانات.

ثانيا: العوامل المحركة لثورة التنقيب في البيانات

يمكن تقسيم العوامل المحركة لثورة التنقيب في البيانات إلى قسمين رئيسيين هما: (5)

نماذج التنقيب في البيانات نوعان: النماذج التنبؤية (*Predictive Data Mining*)

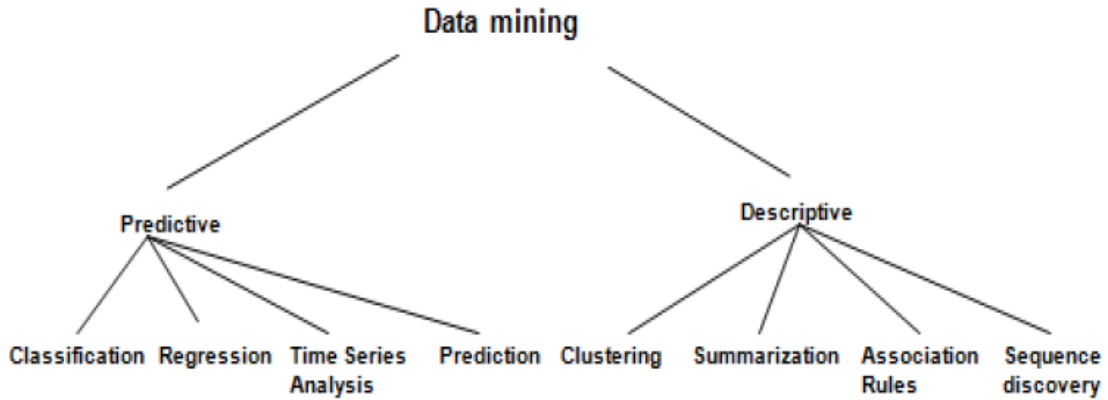
يحاول إيجاد أفضل التنبؤات اعتمادا على المعطيات ويعتمد هذا التنقيب على

استخدام المعلومات القديمة لتوقع ما سيحدث في المستقبل، والنماذج الوصفية

(*Descriptive Data Mining*) تعتمد علي إعادة تنظيم البيانات والتقيب في

أعماقها لاستخراج المؤشرات الموجودة فيها.

شكل رقم 1: أنواع ومهام تقنية تنقيب البيانات



Mining Introductory and Advanced Topics, Prentice Hal 2003

تناول شكل رقم 1: نماذج التنقيب في البيانات وتنقسم إلى نوعان: النماذج التنبؤية (*Predictive*) والنماذج الوصفية (*Descriptives*)، وهناك عدة أدوات للتنقيب في البيانات، نذكر أهمها:

1. السلاسل الزمنية *Time Series Analysis*

السلسلة الزمنية هي عبارة عن قيم ظاهرة من الظواهر في سلسلة تواريخ متلاحقة، أياما أو أشهرها أو سنوات والهدف من وراء تحليل سلسلة زمنية هو دراسة التغيرات التي تكون قد طرأت على الظاهرة التي تمثلها خلال فترة زمنية وتحليل أسبابها ونتائجها أو التنبؤ اعتمادا على فكرة من حوادث الماضي للمستقبل⁽⁶⁾.

الهدف من دراسة السلسلة الزمنية وتحليلها هو تعرف التغيرات التي طرأت على الظاهرة التي تمثلها في مدة من الزمن، ثم تحليل أسبابها ونتائجها وتحديد اتجاهها حتى يمكن استخدامها للتقدير والتنبؤ بالمستقبل، وللتنبؤ بسلوك مسار الاتجاه العام للظاهرة في المستقبل يجب استخدام أحد منحنيات النمو المعبر عنها بعلاقة رياضية أو بنموذج رياضي وباستخدام هذا النموذج الرياضي يمكن تحديد معدل نمو السلسلة الزمنية موضوع الدراسة وتحديد الاتجاه العام لهذه السلسلة وتحديد القيمة التي يمكن أن يصل إليها هذا الاتجاه في المستقبل.

وتحليل السلاسل الزمنية هي تقدير نموذج رياضي يمكنه أن يحاكي تقريبا التدرج التاريخي لتلك الظاهرة بحيث يمكنه أن يقدر بدقة قيم السلسلة الزمنية ويمكن استخدامه بالتنبؤ بقيم مستقبلية لهذه الظاهرة (7).

2. التصنيف (Classification):

يتمثل التصنيف في تفسير أو التنبؤ بخاصية فرد ما من خلال خصائص أخرى. هذه الخاصية هي عموما كيفية (7). ويمكن إنجاز التصنيف بالاعتماد على الأساليب الإحصائية القديمة مثل الانحدار والتحليل التمييزي، أو بالاعتماد على أساليب حديثة نسبيا مثل قوى الارتباط والاستنتاج المستند إلى الحالة والشبكات العصبية.

3. التنبؤ (Prediction):

يشبه التنبؤ التصنيف أو التقدير، ما عدا أن البيانات تصنف على أساس التنبؤ بسلوكها المستقبلي أو تقدير قيمتها المستقبلية (8). حيث أن المتغير التابع المنتبأ به هو متغير كمي. ومن الأدوات التقليدية المستخدمة في التنبؤ نذكر على سبيل المثال: الانحدارات بأنواعها والتحليل التمييزي. أما الأساليب الجديدة فتشتمل على قواعد الارتباط و شجرة القرار والشبكات العصبية والخوارزميات الوراثية.

4. التلخيص (Summarization):

يشير التلخيص إلى أساليب تفتت كتل البيانات الكبيرة إلى مقاييس موجزة، توفر وصفا عاما للمتغيرات وعلاقتها (6). ومن الأمثلة على أساليب التلخيص نذكر: المتوسطات، والمجاميع، والإحصائيات الوصفية التي تتضمن مقاييس النزعة المركزية مثل المتوسط الحسابي والوسيط والمنوال، ومقاييس التشتت مثل الانحراف المعياري.

5. العقودة أو التجزئة (Clustering):

يقتل التجميع العقودي أو التجزئة إلى قطاعات في البحث عن مجموعات متجانسة في مجتمع من الأفراد (9). ويشير التجميع العقودي أو التجزئة إلى قطاعات إلى عملية تشكيل مجموعات أو قطاعات مؤلفة من أفراد أو أصحاب أسر، وذلك بالاستناد إلى معلومات متضمنة في مجاميع من المتغيرات التي تصفهم. والغرض من التجميع العقودي المساعدة على تطوير برامج تسويقية مصممة على مقاسات الزبائن أنفسهم، والتي بالإمكان استخدامها لاستهداف أعضاء لكل قطاع من هذه القطاعات على أمل ترغيبهم في تكرار الشراء أو التحول إلى زبائن موالين (10).

وتتم أساليب التجميع العنقودي غالبا بمساعدة أساليب التحليل العنقودي الإحصائية و الأساليب المستندة إلى شجرة القرار، و الشبكات العصبية و الخوارزميات الوراثية.

6. تحليل الارتباط (*Rule Analysis*):

يتمثل الارتباط في البحث عن علاقات أو ارتباطات موجودة بين عدة خصائص⁽¹¹⁾. و يشير تحليل الارتباط إلى مجموعة من الأساليب التي تستخدم لربط أنماط الشراء عبر القطاعات المتقاطعة أو عبر الوقت. فمثلا يقوم أسلوب تحليل سلة السوق (نوع من أنواع الارتباط) باستخدام المعلومات الكامنة في السلع التي اشتراها المستهلكون فعليا للتنبؤ بالسلع المحتمل شراؤهم إياها إذا ما تم تقديم عروض خاصة لهم أو إذا تم تعريفهم بهذه السلع⁽¹²⁾.

7. الكشف عن التغيرات أو الانحرافات (*Change and deviation detection*)

يرتكز على استكشاف التغيرات المهمة جدا في البيانات من خلال قياسات سابقة أو قيم معيارية.

ثالثا: مراحل عملية التنقيب في البيانات

يمكن تلخيص مراحل و خطوات عملية التنقيب في البيانات كما يلي:⁽¹³⁾

1. فهم طبيعة الأعمال (*Business Understanding*): يعتبر المطلب الأول لاكتشاف المعرفة هو فهم المشاكل و المسائل التي تواجهها الأعمال. و بمعنى آخر، كيف يمكن تحقيق المنفعة الأعظم من التنقيب في البيانات، مما يتطلب وجود صيغة واضحة و محددة لأهداف الأعمال.

2. فهم البيانات (*Data Understanding*): تعتبر مسألة معرفة ماهية وطبيعة البيانات عامل مهم في نجاح عملية التنقيب في البيانات و اكتشاف المعرفة. حيث أن معرفة البيانات بصورة جيدة تعني مساعدة المصممين على استخدام الخوارزميات أو الأدوات المستخدمة للمسائل المحددة بدقة عالية. و هذا يقود إلى تعظيم فرص النجاح بالإضافة إلى رفع الفعالية و الكفاءة لنظام اكتشاف المعرفة. و لا تحتاج عملية التنقيب في البيانات إلى تجميع البيانات في مستودع البيانات، أما إذا كان مستودع البيانات موجود في المنظمة، فمن الأفضل عدم احتكار المستودع بشكل مباشر لغرض التنقيب في البيانات. و يمكن تلخيص الخطوات الضرورية لعملية فهم البيانات كالتالي:

- **تجميع البيانات (Data Collection):** و هي الخطوة الموجهة نحو تحديد مصدر البيانات في الدراسة بما في ذلك استخدام البيانات العامة الخارجية مثل الضرائب و غيرها.
 - **توصيف البيانات (Data Description):** و هي الخطوة التي تركز على توصيف محتويات الملف الواحد من الملفات أو الجداول.
 - **جودة البيانات و تحقيقها (Data Quality and Verification):** هذه الخطوة تحدد ما إذا كان تقليل أو إهمال بعض البيانات غير الضرورية أو كونها رديئة الجودة و قد لا تنفع في الدراسة. لأن النموذج الجيد يحتاج إلى بيانات جيدة مما يتوجب أن تكون البيانات صحيحة و ذات مضمون دقيق.
 - **التحليل الاسترشادي للبيانات (Exploratory Analysis of Data):** تستخدم الأساليب مثل الإظهار المرئي أو التصور أو عملية التحليل المباشر (OLAP) التي تؤدي إلى إجراء التحليل الأولي للبيانات. و تعتبر هذه الخطوة مهمة و ضرورية لأنها تركز على تطوير الفرضيات المتعلقة بالمشكلة قيد الدراسة.
3. **تهيئة البيانات (Data Preparation):** و تشمل الخطوات التالية:
- **الاختيار (Selection)** و تعني اختيار المتغيرات المتوقعة و حجم العينة.
 - **صياغة المتغيرات و تحويلها (Construction and Transformation Variables)** حيث يجب دائما أن تصاغ المتغيرات الجديدة لبناء النماذج الفعالة.
 - **تكامل البيانات (Data Integration):** حيث أن مجاميع البيانات في دراسة التنقيب عن البيانات من الممكن خزنها في قواعد بيانات متعددة الأغراض التي تكون بحاجة إلى توحيدها في قاعدة بيانات واحدة.
 - **تصميم و تنسيق البيانات (Data Formatting)** حيث تتعلق هذه الخطوة في إعادة ترتيب حقول البيانات كما يتطلب في نموذج التنقيب في البيانات.
4. **صياغة نماذج الحل و ثبوتها (Model Building and Validation)**

إن بناء و صياغة نموذج الحل السليم و الدقيق يتم من خلال عملية الخطأ و الصواب، حيث كثيرا ما تحتاج مثل هذه العملية إلى مساعدة المختصين في التنقيب عن البيانات بهدف اختبار و فحص مختلف البدائل للحصول على أفضل نموذج لحل المشكلة قيد الدراسة.

5. التقييم و تحليل نتائج النموذج (Evaluation and Interpretation):

حالما يتم صياغة النموذج و التحقق من ثباته و صدقه، تجري مباشرة عملية التحقق من ثبات حزمة البيانات التي يتم تغذيتها بواسطة النموذج. وبما أن نتائج هذه البيانات معروفة، لذا فان النتائج المتوقعة تقارن مع النتائج الفعلية في ثبات حزمة البيانات قيد التشغيل. و تؤدي هذه المقارنة أو المفاضلة إلى التحقق من دقة النموذج.

6. نشر و توزيع النموذج (Model Deployment):

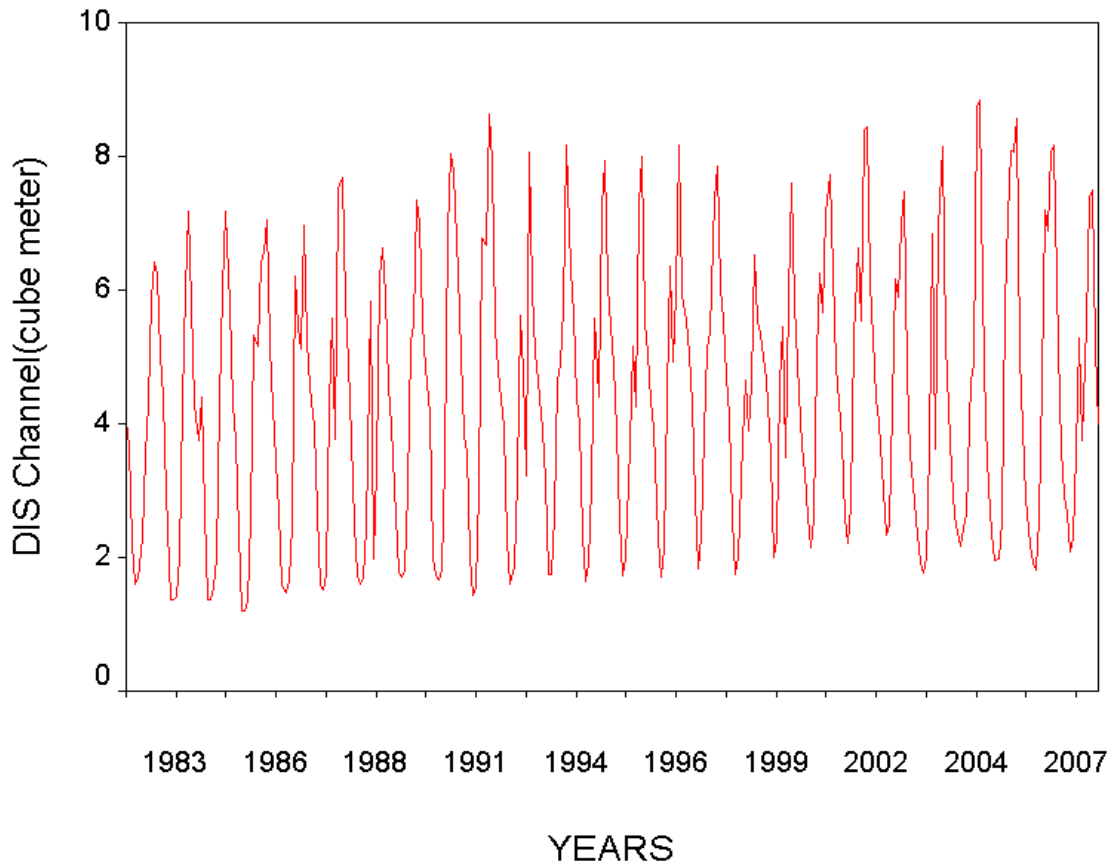
حيث تشتمل هذه الخطوة على نشر و توزيع النموذج داخل المنظمة لمساعدة عملية صنع القرار. و أن النموذج الصالح يجب أيضا أن يحقق الرضا لدى المستخدمين طالما أن اختيار النموذج لا بد أن يتم من خلال الدراسة الاسترشادية أو نموذج مصغر من الدراسة الشاملة.

الإطار التطبيقي:

تم الحصول على بيانات الدراسة من إدارة خزان خشم القرية وهي عبارة عن المناسيب السنوية مقاسه بالمتر المكعب، تم تطبيق البيانات للسلسلة الزمنية (هي مجموعة من المشاهدات الخاصة بظاهرة فاقد التربة الرئيسية للفترة 1981-2007م) والنتائج المتحصل عليها باستخدام تقنية تنقيب البيانات أسلوب السلاسل الزمنية كما تمر بالمرحل كما بالشكل 4-6 تمر بالمرحل التالية:

أولا التشخيص

ولتوضيح الأسس السابقة كما بالإطار النظري تم تطبيق متغير فاقد التربة الرئيسية والذي يمثل بياناته سلسلة زمنية وهي عبارة عن مناسيب قراءات يومية للمياه للخزان لابد من رسم السلسلة بيانيا باستخدام شكل السلسلة الزمنية عبر الزمن كما في الشكل التالي.



شكل رقم: 1-Error! No text of specified style in document. رسم بيانات السلسلة الزمنية لفاقد التربة

يتضح من خلال الشكل رقم 4-15 البياني للسلسلة متغير فاقد التزعة عدم وجود اتجاه عام و هذا يعني إن السلسلة مستقرة في وسطها و كذلك نلاحظ عدم وجود تغيرات منتظمة في السلسلة. من تطبيق أدوات التشخيص لمعرفة نوع النموذج تحصلنا علي نموذج من الدرجة الأولي وعند تشخيصية اتضح انه النموذج المناسب.

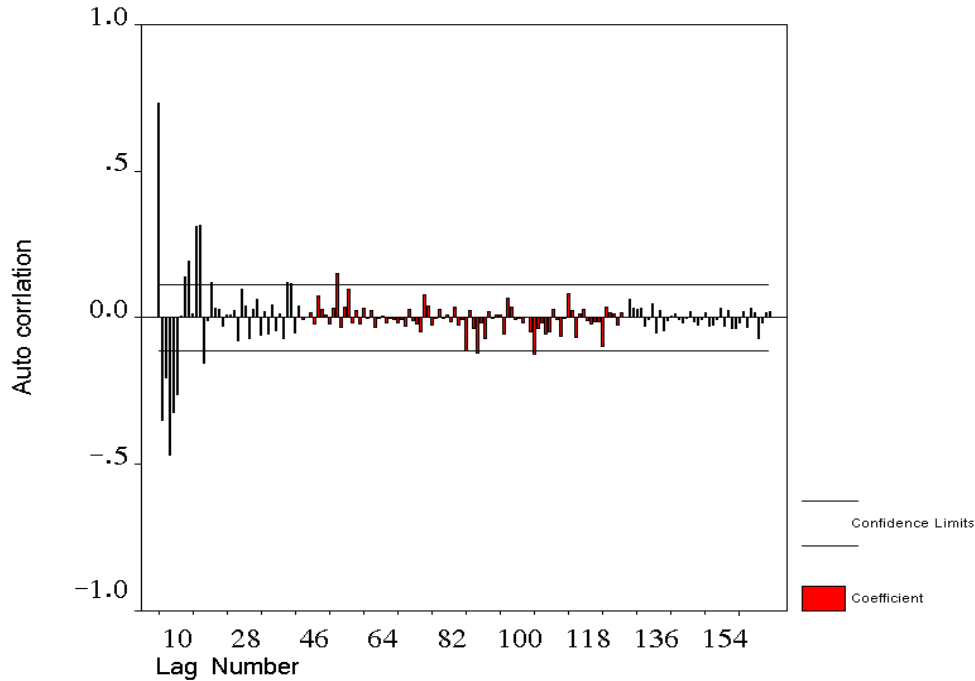
ثانيا تقدير معلمات النموذج

ولتقدير معلمات النموذج لابد من معرفة النموذج وتحديد رتبته والتأكد من ثباته واستقرار التباين كما موضحة كالتالي:

جدول رقم: 1-Error! No text of specified style in document: بعض المقاييس

الإحصائية لمنسوب فاقد التزعة

فاقد التزعة	Std.deviation	Mean	Maximum	Minimum
	1.986	4.348	8.848	1.195



شكل رقم: 2-Error! No text of specified style in document: رسم بيانات السلسلة

الزمنية لمنسوب فاقد التزعة

يلاحظ من الرسم شكل رقم 4-16 إن المتغير فاقد التربة الرئيسية إن جميع معاملات الارتباط الجزئي تقع داخل فترة الثقة، وهذا يدل علي إن السلسلة الزمنية ساكنة (البيانات عشوائية).

ثالثا فحص مدي ملائمة النموذج (فاقد الترتة الرئيسية)

لفحص مدي ملائمة النموذج نقوم بإجراء فحص البيانات الوصفية لمنسوب الفاقد واختبار المحاكاة لنموذج الانحدار الذاتي لسلسلة لمنسوب فاقد الترتة كما يلي:

جدول رقم: Error! No text of specified style in document. 2- نموذج الانحدار

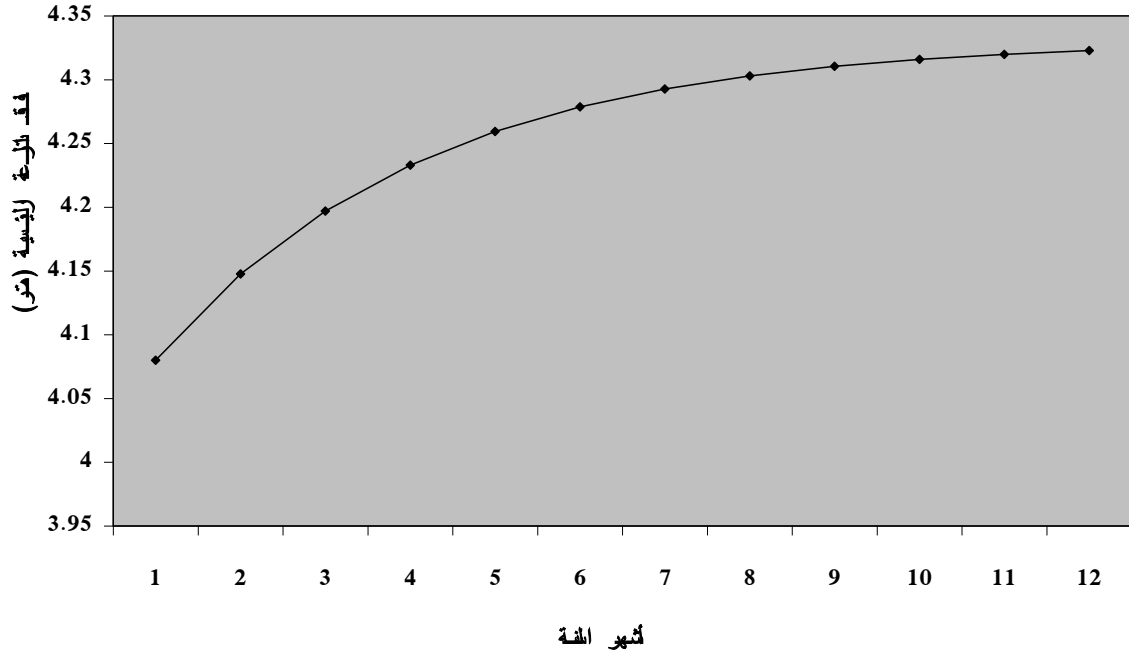
الذاتي من الدرجة الأولى لسلسلة فاقد الترتة

فاقد الترتة	المقدرات	الخطأ المعياري	اختبار -ت	القيمة الاحتمالية
AR1	.7307	.03858	18.939	.0000
constant	4.3423	.28213	15.391	.0000

ثالثا لفحص مدي الملائمة واختبار المحاكاه نجد من الجدول رقم 4-8 أن القيم الاحتمالية المقدرات للنموذج AR2 تساوي أصفار و هذا يعني أن جميع المقدرات معنوية و لها تأثير ذو دلالة إحصائية.

رابعاً التنبؤ لمنسوب السلسلة الزمنية لمنسوب خلف التربة الرئيسية

التنبؤ للعام 2007م لمنسوب لفاقد التربة الرئيسية



نلاحظ إن القيم المتوقعة متزايدة لفاقد التربة مما يؤثر سلبي علي ري المساحات وبالتالي ينعكس علي إنتاج المحاصيل في المشروع.

الخلاصة:

إن إدماج مزايا تكنولوجيا المعلومات مع الطرق الإحصائية و الخوارزميات قاد إلى توفر الإمكانيات اللازمة للتنبؤ بالسلوك المستقبلي و من ثم وضع الحلول المناسبة للمشكلات قبل وقوعها في حال إمكان حدوثها، أو من باب التنبؤ بهدف التطوير و التحديث بشكل عام في شتى المجالات كل ذلك باستخدام تقنيات التنقيب في البيانات التي تشكل مرحلة من مسار أشمل هو استكشاف المعرفة في قواعد البيانات و التي أصبحت من الهموم الكبيرة التي تقع على عاتق الدول بكافة مؤسساتها بشكل عام.

المراجع:

1. بشير عباس، العلاق، الإدارة الرقمية: المجالات و التطبيقات، مركز الإمارات للدراسات و البحوث الإستراتيجية، ابوظبي، 2005.
2. Bazsalica M., Naim P., Data mining pour le Web, éd. Eyrolles, Paris, 2001.
3. عبد الستار العلي، عامر إبراهيم قنديلجي، غسان العمري، المدخل إلى إدارة المعرفة، دار المسيرة للنشر و التوزيع و الطباعة، الطبعة الأولى، عمان، 2006.
4. Hand d., Mannila H., Smyth R., Principles of Data Mining, MIT Press, London.
5. بشير عباس، العلاق، مرجع سابق.
6. بشير عباس، العلاق، مرجع سابق.
7. د.نورة عبد الرحمن اليوسف: محاضرات الاقتصاد القياسي، جامعة الملك سعود، كلية العلوم الإدارية قسم الاقتصاد.
8. Bazsalica M., Naim P.
9. Berry J. A. M., Linoff G. S., Data Mining Techniques For Marketing, Sales, and Customer Relationship Management, 2° ed., Wiley Publishing, INC, Indianapolis.
10. Bazsalica M., Naim P.
11. بشير عباس العلاق، مرجع سابق.
12. Bazsalica M., Naim P.
13. بشير عباس العلاق، مرجع سابق.
14. عبد الستار العلي، عامر إبراهيم قنديلجي، غسان العمري، مرجع سابق.
15. عبد الستار العلي، عامر إبراهيم قنديلجي، غسان العمري، مرجع سابق.
16. د.طيبار أحسن، بشلابي عمار، التنقيب في البيانات واتخاذ القرارات.