بسم الله الرحمن الرحيم

جامعة آل-البيت

عمادة البحث العلمي والدراسات العليا

كلية الأمير الحسين بن عبد الله لتكنولوجيا المعلومات

قسم الحاسوب

رسالة ماجستير بعنوان

التنقيب عن البيانات باستخدام الخوارزميات الجينية

Data Mining Using Genetic Algorithms

إعداد

مازن سالم حمد الزيود

٠٤٢٠٩٠١٠٠٣

المشرف

الدكتور جهاد قبيل ألخالدي

٢٠٠٨

المنارة للاستشارات

www.manaraa.com

التنقيب عن البيانات باستخدام الخوارزميات الجينية

Data Mining Using Genetic Algorithms

إعداد

مازن سالم حمد الزيود

٠٤٢٠٩٠١٠٠٣

المشرف

الدكتور جهاد قبيل ألخالدي

أعضاء لجنة المناقشة		التوقيع
• د.جهاد قبيل الخالدي	مشرفا/ رئيسا	_____
• أ.د عدنان الصمادي	عضو	_____
• د. مأمون ربابعة	عضو	_____
• د. غسان كنعان	عضو	_____

قدمت هذه الرسالة استكمالا لمتطلبات الحصول على درجة الماجستير في علم الحاسوب في كلية الأمير الحسين بن عبد الله لتكنولوجيا المعلومات في جامعة آل-البيت.

نوقشت وأوصي بإجازتها بتاريخ / / ٢٠٠م

المنارة للاستشارات

www.manaraa.com

الإهداء

إلى الذي علمني معنى الحياة وعنوان الكفاح..... والدي

إلى ملكة طفولتي وأميرة شبابي رمز العطاء ..... والدتي.

إلى رونق حياتي وسر سعادتي ..... إخواني

المنارة للاستشارات
www.manaraa.com

كلمة شكر

أتقدم بخالص شكري وعظيم امتناني إلى جامعة آل-البيت ممثلة برئيسها وأعضاء هيئة التدريس وأخص بالشكر **الدكتور جهاد قبيل الخالدي** الذي كان له اكبر الأثر في إخراج هذه الرسالة آملا أن أكون قد حققت الهدف المرجو منها.وكما أتقدم بالشكر والعرفان لأعضاء لجنة المناقشة على الوقت والجهد الذي بذلوه لقراءة ومناقشة هذه الرسالة والى كل من قدم لي يد العون والمساعدة.

مازن سالم حمد الزيود
٠٤٢٠٩٠١٠٠٣

المنارة للاستشارات
www.manaraa.com

المحتويات

الموضـــــــــوع	رقم الصفحة
الغلاف الداخلي	أ
الإهداء	ب
الشكر	ج
المحتويات	د
فهرس الاشكال	ز
فهرس الجداول	ط
الملخص	ي
الفصل الأول	
الفصل التمهيدي	
١ تعريف المشكلة والاهداف	٢
١-١ التعريف	٢
٢- أهمية الدراسة	٣
٣- اهداف الدراسة	٤
الفصل الثاني	
الدراسات السابقة	
٢-١ استخدام الخوارزميات الجينية في تحسين الانتاجية	٦
٢-٢ استخدام نظام خبير للوصول الى قواعد التنقيب عن البيانات	٦
٢-٣ استخدام الخوارزميات الجينية لاستخلاص قواعد المعرفة	٧
٢-٤ تقنيات الحساب العصري	٧
٢-٥ دراسة حول الخوارزميات الجينية	٨
٢-٥-١ الخوارزميات الجينية ليست محسن لغايات عامة	٨
٢-٥-٢ الخوارزميات الجينية مختلفة جدا	٨
٢-٥-٣ بماذا يمكن ان تكون الخوارزميات الجينية مناسبة	٩
الفصل الثالث	
الخوارزميات الجينية	
٣-١ مقدمة عن الخوارزميات التطورية	١١
٣-٢ ميزات الخوارزميات التطورية	١١
٣-٣ معاملات الخوارزميات التطورية	١٢
٣-٣-١ الاختيار	١٢
٣-٣-٢ التهجين	١٢
٣-٣-٣ الطفرة	١٣
٣-٣-٤ اعادة الدمج	١٣
٣-٣-٥ نماذج المجتمع – بالتوازي مع تنفيذ الخوارزميات التطورية	١٤
٣-٣-٦ تطبيق استراتيجيات متعددة\ مختلفة والمنافسة بين المجتمعات الفرعية	١٤
٣-٤ مقدمة عن الخوارزميات الجينية	١٥

المنارة للاستشارات
www.manaraa.com

٣-٥ تعريف بالخوارزميات الجينية	١٥
٣-٦ نقاط رئيسية	١٦
٣-٧ كيفية عمل الخوارزميات الجينية	١٦
٣-٨ الخوارزميات الجينية كـ(pseudo code)	١٦
٣-٩ مشكلات تخص الخوارزميات الجينية	١٧
٣-١٠ مشكلات حجم المجتمع	١٨
٣-١١ دالة الياقة	١٨
٣-١٢ معاملات الخوارزميات الجينية	١٨
٣-١٢-١ الطفرة	١٨
٣-١٢-٢ معامل التداخل	١٩
٣-١٢-٣ القلب واعادة التجميع	٢٠
٣-١٣ تطبيقات الخوارزميات الجينية	٢١
٣-١٤ حقول من تعلم الآلة	٢١
٣-١٤-١ تجميع البيانات	٢١
٣-١٤-٢ التحقيق المتقاطع	٢٢
٣-١٥ التنقيب عن البيانات	٢٤
٣-١٦ ما هو التنقيب عن البيانات	٢٤
٣-١٧ التنقيب عن البيانات من الناحية التاريخية	٢٥
٣-١٨ ظهور التنقيب عن البيانات	٢٥
٣-١٩ التنقيب في سجلات مركز مواد تموينية	٢٥
٣-٢٠ عملية التنقيب عن البيانات	٢٦
٣-٢١ خطوات التنقيب عن البيانات	٢٧
٣-٢٢ الجهات المهتمة بالتنقيب عن البيانات	٢٨
٣-٢٣ كيف يعمل التنقيب عن البيانات	٢٩
٣-٢٤ العناصر الرئيسة في التنقيب عن البيانات	٣٠
٣-٢٥ اكتشاف المعرفة في قواعد البيانات	٣٠
٣-٢٥-١ الخطوة المعنية باكتشاف المعرفة في قواعد البيانات بشكل عام	٣١
٣-٢٦ تطبيقات على التنقيب عن البيانات	٣١
الفصل الرابع	
النظام المقترح	
٤-١ مقدمة	٣٤
٤-٢ الطريقة	٣٤
٤-٢-١ طريقة البحث المعتمدة	٣٤
٤-٣ تحليل النظام المقترح	٣٨
٤-٣-١ دراسة المتطلبات	٣٩
٤-٣-٢ المستخدمون للنظام	٣٩
٤-٣-٢-١ نموذج حالة الاستخدام	٣٩
٤-٣-٣ التعريف بالمتطلبات العملية	٤٠
٤-٣-٤ مواصفات العمليات	٤٠
٤-٣-٥ تصميم النظام	٤٦
٤-٣-٥-١ مخططات تسلسل الاحداث والتعاون	٤٦
٤-٣-٥-٢ مخططات الاحداث	٥٣
٤-٤ التطبيق	٥٦
٤-٤-١ كيفية استخدام هذا البرنامج	٥٦

المنارة للاستشارات
www.manaraa.com

الفصل الخامس	
النتائج	
٥-١ مقدمة	٦٢
٥-٢ مزايا النظام المقترح	٦٢
٥-٣ نتائج النظام	٦٢
٥-٣-١ نتائج عملية تشغيل النظام	٦٢
٥-٣-٢ نتائج عملية التزاوج الاختياري	٦٦
٥-٤ نتائج البحث والعمل المستقبلي	٧١
المصادر والمراجع	٧٢
الملخص باللغة الانجليزية	٨٤

المنارة للاستشارات
www.manaraa.com

فهرس الأشكال

الشكل	الصفحة
الشكل (٣ – ١) تطبيق عملية التداخل Crossover	٢٠
الشكل (٣ – ٢) القلب واعادة التجميع	٢١
الشكل (٤-١) مراحل التطوير التكراري والتزايدي	٣٥
الشكل (٤ – ٢) المعلومات المجدولة لسلع المركز التمويني	٣٧
الشكل (٤ – ٣) نموذج حالة الاستخدام	٣٩
الشكل (٤-٤) تسلسل الاحداث لعملية التشغيل	٤٦
الشكل (٤-٥) التعاون في العمليات لعملية التشغيل	٤٧
الشكل (٤-٦) تسلسل الاحداث لعملية الاغلاق	٤٨
الشكل (٤-٧) التعاون في العمليات لعملية الاغلاق	٤٨
الشكل (٤-٨) تسلسل الاحداث لعملية المسح	٤٩
الشكل (٤-٩) التعاون في العمليات لعملية المسح	٤٩
الشكل (٤-١٠) تسلسل الاحداث لعملية التزاوج الاختياري	٥٠
الشكل (٤-١١) التعاون في العمليات لعملية التزاوج الاختياري	٥١
الشكل (٤-١٢) تسلسل الاحداث لعملية تحديث قاعدة البيانات	٥١
الشكل (٤-١٣) التعاون في العمليات لعملية تحديث قاعدة البيانات	٥٢
الشكل (٤-١٤) الاحداث لعملية التشغيل	٥٣
الشكل (٤-١٥) الاحداث لعملية الاغلاق	٥٤
الشكل (٤-١٦) الاحداث لعملية المسح	٥٤
الشكل (٤-١٧) الاحداث لعملية التزاوج الاختياري	٥٥
الشكل (٤-١٨) الاحداث لعملية تحديث قاعدة البيانات	٥٦
الشكل (٤-١٩) الواجهة الرئيسية للنظام	٥٧
الشكل(٤-٢٠) ناتج تنفيذ عملية التشغيل	٥٨

المنارة للاستشارات
www.manaraa.com

الشكل (٤- ٢١) ناتج عملية تجديد قاعدة البيانات	٦٠
الشكل(٥-١) نتائج تشغيل النظام	٦٣
الشكل(٥-٢) نتائج تشغيل النظام	٦٣
الشكل(٥-٣) نتائج تشغيل النظام	٦٤
الشكل(٥-٤) نتائج تشغيل النظام	٦٤
الشكل(٥-٥) نتائج تشغيل النظام	٦٥
الشكل(٥-٦) نتائج تشغيل النظام	٦٥
الشكل(٥-٧) نتائج عملية التزاوج الاختياري	٦٧
الشكل(٥-٨) نتائج عملية التزاوج الاختياري	٦٧
الشكل(٥-٩) نتائج عملية التزاوج الاختياري	٦٨
الشكل(٥-١٠) نتائج عملية التزاوج الاختياري	٦٩
الشكل(٥-١١) نتائج عملية التزاوج الاختياري	٦٩
الشكل(٥-١٢) مقارنة بين نتائج للنظام المقترح والتنفيذ اليدوي	٧٠

المنارة للاستشارات
www.manaraa.com

فهرس الجداول

الجدول	الصفحة
الجدول (٤-١) عملية تشغيل التنقيب عن البيانات	٤١
الجدول (٤ – ٢) عملية مسح المعلومات	٤٢
الجدول (٤ – ٣) عملية اغلاق النظام	٤٣
الجدول (٤ – ٤) عملية التزاوج الاختياري	٤٤
الجدول (٤ – ٥) عملية تجديد قاعدة البيانات	٤٥
الجدول(٥-١) نتائج التجارب للنظام المقترح والنظام اليدوي	٧٠

المنارة للاستشارات
www.manaraa.com

الملخص

يعد التنقيب عن البيانات من أكثر حقول تكنولوجيا المعلومات أهمية نظراً لتزايد البيانات بشكل هائل جداً وصعوبة استخراج معلومات مهمة من هذا الكم، والتنقيب عن البيانات أصبح مستخدماً في الكثير من المجالات وهذه المجالات تتضمن شركات البطاقات الائتمانية والبنوك و التسويق والصناعة وشركات الهاتف وغيرها الكثير. وبشكل عام تنقسم طرق التنقيب عن البيانات إلى قسمين، الأول هو استخدام الإحصاءات، والثاني والذي يعد أحد أهم حقول الذكاء الاصطناعي هو حقل تعلم الآلة(Siddhartha, 2002).

ومن خلال هذه الدراسة سنقوم باستخدام الخوارزميات الجينية في عملية استخراج البيانات والتنقيب عنها من قاعدة بيانات تمثل مبيعات لمركز تمويني. وهنالك عدد من المزايا التي تجعل من الخوارزميات الجينية مناسبة لحل المشكلات الصعبة والكبيرة، حيث أن الخوارزميات الجينية صممّت خصيصاً لتقدم أفضل أداء مع المتغيرات غير الخطية ومع البيانات التي تحتوي على أخطاء بالإضافة إلى مقدرتها الكبيرة على البحث في مجموعة كبيرة من الأمثلة المقيّدة بقواعد معينة بحثاً عن الحل الأمثل.(Mustafa, 2003)،(Binti Abdullah,et.al, 2003).

يتميز النظام المقترح ببناء نظام جديد للتنقيب عن البيانات، وقد قمنا من خلال هذه الدراسة بتطوير عمليه تداخل بحيث تأخذ أفضل عناصر من كروموسومين، ويراعى فيها ان يتمكن المستخدم من التحكم في طول الكروموسوم وعملية التزاوج بين أشهر يحددها المستخدم، وإيجاد عدد من الكروموسومات من خلال عملية التزاوج وظهور كروموسومات جديدة. ونقدم خوارزميه جينية مبنية لغايات تحسين اتخاذ القرارات.

المنارة للاستشارات
www.manaraa.com

الفصل الأول

المقدمة

المنارة للاستشارات
www.manaraa.com

١. تعريف المشكلة والأهداف:

١.١. التعريف:

يعد التنقيب عن البيانات من أكثر حقول تكنولوجيا المعلومات أهمية نظراً لتزايد البيانات بشكل هائل جداً وصعوبة استخراج معلومات مهمة من هذا الكم، والتنقيب عن البيانات أصبح مستخدماً في الكثير من المجالات وهذه المجالات تتضمن شركات البطاقات الائتمانية والبنوك و التسويق والصناعة وشركات الهاتف وغيرها الكثير (Siddhartha, 2002).

وبشكل عام تنقسم طرق التنقيب عن البيانات إلى قسمين، الأول هو استخدام الإحصاءات، والثاني والذي يعد أحد أهم حقول الذكاء الاصطناعي هو حقل تعلم الآلة ويفترض استخدام مجموعة من البيانات التدريبية والتي من خلالها يتعلم نظام التنقيب عن البيانات كيفية استخراج المعلومات المهمة (Siddhartha, 2002).

وتعلم الآلة هو من حقول الذكاء الاصطناعي الأكثر أهمية حيث أنه هو السبب الرئيسي لظهور هذا المجال.ويتضمن هذا الحقل مجموعة من الأجزاء ومنها الشبكات العصبية والخوارزميات الجينية. وفي الشبكات العصبية يتم الربط بين مجموعة من الطبقات وفي كل طبقة مجموعة من عصيبات تتصل مع بعضها البعض وتتم عملية التعلم من خلال خوارزمية تقوم على تغيير الأوزان بين هذه العصيبات (Lugar, 2002).

أما الخوارزميات الجينية فتعتمد على مجموعة من الكروموسومات التي تحمل مقدار من اللياقة يحدد مدى كفاءتها وتقوم بالتزاوج لإخراج كروموسومات أفضل (Lugar, 2002).

المنارة للاستشارات
www.manaraa.com

ومن خلال هذه الدراسة سنقوم باستخدام الخوارزميات الجينية في عملية استخراج البيانات والتنقيب عنها من قاعدة بيانات تمثل مبيعات لمركز تمويني. وهنالك عدد من المزايا التي تجعل من الخوارزميات الجينية مناسبة لحل المشكلات الصعبة والكبيرة، حيث أن الخوارزميات الجينية صممّت خصيصاً لتقدم أفضل أداء مع المتغيرات غير الخطية ومع البيانات التي تحتوي على أخطاء بالإضافة إلى مقدرتها الكبيرة على البحث في مجموعة كبيرة من الأمثلة المقيّدة بقواعد معينة بحثاً عن الحل الأمثل كما أن لها المقدرة على تحويل القواعد إلى عبارات رقمية.(Mustafa, 2003)،(Binti Abdullah,et.al, 2003).

٢. أهمية الدراسة:

تكمن اهمية هذه الدراسة في لتعريف بالتنقيب عن البيانات، ماهيته، أهميته، والطرق المستخدمة في تطبيقه وسيقوم البحث بتوضيح مجمل زوايا هذا الحقل من حقول تكنولوجيا المعلومات. تقوم هذه الدراسة أيضاً بالتعريف بالخوارزميات الجينية، ماهيتها، أهميتها، طرقها وكيفية استخدامها في الحقول الأخرى.

ومن خلال هذه الدراسة سنعرف ونشرح الجانب الرئيسي منها، ألا وهو استخدام الخوارزميات الجينية في التنقيب عن البيانات وتطبيقها على أرض الواقع من خلال برنامج محاكاة (Simulation) يقوم على تحسين المبيعات والتنظيم في المراكز التموينية.

المنارة للاستشارات
www.manaraa.com

٣. اهداف الدراسة:

ستقدم هذه الدراسة مساهمات في مجال استخدام الخوارزميات الجينية في التنقيب عن البيانات وذلك من خلال ما يلي:

أولاً: سنقوم بتطبيق الخوارزميات الجينية على قاعدة بيانات تم استحداثها لهذه الغاية.

ثانياً: معرفة مدى ملائمة وكفاءة الخوارزميات الجينية في التنقيب عن البيانات.

ثالثا: بناء نظام جديد ذو فاعلية في التنقيب عن البيانات يراعى فيه أن يتمكن المستخدم من التحكم في طول الكروموسوم وعمل تزاوج بين أشهر معينة في السنة، بالإضافة إلى إيجاد كروموسومات جديدة من خلال إيجاد عدد من الكروموسومات العشوائية بالإضافة إلى عمليات تغيير(Mutation).

المنارة للاستشارات
www.manaraa.com

الفصل الثاني

الدراسات السابقة

المنارة للاستشارات
www.manaraa.com

يتم في هذا الفصل عرض بعض الدراسات السابقة، التي بعضها أستخدم الخوارزميات الجينية في اكتشاف قواعد منطقية وتحسين الإنتاجية، وغيرها أتجه إلى حقول مثل تعلم الآلة ومن هذه الدراسات ما يلي:

٢.١. استخدام الخوارزميات الجينية في تحسين الانتاجية(Optimization of Production Systems using Genetic Algorithms)

تتحدث هذه الدراسة عن كيفية استخدام الخوارزميات الجينية في تحسين الإنتاجية حيث تجد ترتيب لعناصر شرطية(Condition Elements) CEs حسب قواعد نظام الإنتاج(Production System) PS مما يؤدي إلى نظام أنتاج على كفاءة عالية بالنسبة للوقت (Mustafa, 2003).

ومن مساوئ هذه الدراسة هي أمكانيه عدم مطابقة العناصر الشرطية لنظام الإنتاج نتيجة لطفره جينية او تداخل غير مرغوب وقد قمنا من خلال هذه الدراسة بتطوير عمليه تداخل بحيث تأخذ أفضل عناصر من كروموسومين.

٢.٢. استخدام نظام خبير للوصول الى قواعد التنقيب عن البيانات(Accessing Data Mining Rules through Expert Systems)

في هذه الدراسة تم استخدام نظام خبير للوصول إلى قواعد التنقيب عن البيانات عن طريق أعطاء النظام الخبير هذه القواعد ليقوم بتحليلها بشكل متسلسل، وكيف أن مختلف أنواع قواعد التنقيب عن البيانات يمكن تحويلها إلى معرفة بالنسبة إلى النظام الخبير (Basilis, 2002).

المنارة للاستشارات
www.manaraa.com

عدم مقدرة النظام الخبير على التأقلم مع انواع مختلفه من البيانات وحصره في حيز محدد اما الخوارزميات الجينيه فلها القابليه لذلك.

٢.٣. استخدام الخوارزميات الجينية لاستخلاص قواعد المعرفة(GAs Rule for Knowledge Discovery)

هذه الدراسة تقدم أسلوب جديد لاستخلاص القواعد المهيكلة واكتشاف المعرفة من خلال استخدام الخوارزميات الجينية ويعتمد على استخدام تعلم الرموز لاستنتاج المستقبل كمعالجة مسبقة ومعالجة الرموز الجزئية لاستنتاج القواعد كمعالجة لاحقة (Binti Abdullah,et.al, 2003). يقدم هذا البحث امكانيه توقع المستقبل ولكن لا يقدم امكانيه تحسينه وفي البحث الحالي سنقدم طريقه لتطوير الانتاجيه.

٢.٤. تقنيات الحساب العصري(Evolutionary Computation for database Marketing)

هذه الدراسة تستخدم تقنيات الحساب العصري (Evolutionary Computing) EC مثل الخوارزميات الجينية وبرامجها في التنقيب عن البيانات وبشكل خاص لحل مشكلة التسويق بإستخدام قواعد البيانات، بداية بإعطاء مقدمة موجزة عن البحث باستخدام الجينات (Genetic Search) ومن ثم تقدم لنا نظرة سطحية عن مجالين يفيد بهما استخدام (EC) ألا وهما:

أولاً: تطوير نماذج محسنة تركز على أعماق محددة للبحث فيها.

ثانياً: نماذج تحسن نفسها بشكل متزامن لتحقق عدد من الأهداف (Siddhartha, 2002).

المنارة للاستشارات
www.manaraa.com

يرتكز هذا البحث على مجال التسويق باستخدام البحث الجيني اما البحث الحالي فيطور خوارزميه جينية لتطوير الأرباح بشكل مباشر.

٢.٥. دراسة حول الخوارزميات الجينية(What are Genetic Algorithms good at?)

توضح هذه الدراسة مزايا وإيجابيات الخوارزميات الجينية

٢.٥.١. الخوارزميات الجينية ليست محسن لغايات عامة:

وفي هذا الجزء تبين الدراسة أن ال GAs لا تستخدم بشكل عام بل تبنى لحالات خاصة، حيث أنها تعتمد على قيم حقيقة أعتماداً على النظام والمشكلة التي تواجهه ولهذا السبب تختلف الخوارزمية من حالة إلى حالة. (Peter, 1999)

٢.٥.٢. الخوارزميات الجينية مختلفة جداً:

إن الخوارزميات الجينية تختلف تبعاً للحالة التي يتم معالجتها ولا يمكن تطبيقها على حالات كثيرة مختلفة وحتى تكلفة بناء نظام يستخدمها تختلف تبعاً للخوارزمية المستخدمة في إنتاج الكروموسومات وقياس معدل اللياقة(fitness) لها وإيجاد الأكثر كفاءة (Peter, 1999).

المنارة للاستشارات
www.manaraa.com

٢.٥.٣. بماذا يمكن أن تكون الخوارزميات الجينية مناسبة:

وفي هذا الجزء تتحدث الدراسة عن المجالات التي يمكن أن نستخدم الخوارزميات الجينية فيها وتلقي النظر على بعض هذه المجالات (Peter, 1999).

هذا البحث هو بحث نظري أكثر منه علمياً تطبيقياً حيث يقدم معلومات عن الخوارزميات الجينية اما في بحثنا هذا فنقدم خوارزميه جينية مبنيه لغايات تحسين اتخاذ القرارات.

المنارة للاستشارات
www.manaraa.com

الفصل الثالث

الخوارزميات الجينية

المنارة للاستشارات
www.manaraa.com

٣,١ مقدمة عن الخوارزميات التطورية:

الخوارزميات التطورية Evolutionary Algorithms هي برامج كمبيوترية تحاكي آلية التطور البيولوجي، من تكاثر و تطفر و انتخاب طبيعي، بهدف الوصول إلى حلول مثلى للمشاكل التقنية (Ashlock, 2008).

ان الخوارزميات العادية هي خوارزميات تعتمد على تلقين خطوات متتالية لحل المشكلة، أي أنها خوارزميات تحليلية تفترض معرفة المبرمج مسبقا طريقة حل المشكلة، ولكن في الخوارزميات التطورية لا يكون هناك تصور لطريقة حل المشكلة، كل ما نملكه هو الهدف الذي علينا الوصول له ويقوم البرنامج بتطوير الحل عبر توليد أجيال من الحلول و تكثيرها و الانتخاب فيما بينها كما هو حال التطور البيولوجي (Back, 1996).

كان علماء الاحياء هم أول من بدأ باستخدام الخوارزميات التطورية في محاولتهم فحص بعض العوامل و المتغيرات في نظرية التطور فقاموا ببناء برامج كمبيوتر تحاكي آلية التطور ولم يكن في تصورهم أن هذه البرامج ستجد لها تطبيقات في مجالات بعيدة عن علم الأحياء، مثل هندسة الطيران وصناعة البترول وعلوم الصيدلة والطب. تستخدم الخوارزميات التطورية لحل العديد من المشاكل التقنية في الأبحاث العلمية والصناعة (Back, 1996).

٣،٢ ميزات الخوارزميات التطورية

١. خوارزمية البحث التطوري تبحث في مجموعة من النقاط بشكل متوازي، وليس فقط نقطة واحدة.
٢. خوارزمية البحث التطوري لا تحتاج إلى مشتقات المعلومات أو غيرها من المعارف المساعدة، فقط وظيفة الهدف Objective Function وما يقابلها من مستويات اللياقة Fitness للتأثير على اتجاهات البحث.
٣. خوارزميات البحث التطورية تستخدم قواعد الإنتقال الإحتمالية لا القطعية منها.

المنارة للاستشارات
www.manaraa.com

٤. الخوارزميات التطورية بصفة عامة أكثر وضوحا للتطبيق، لأنها لا وجود لأي قيود لتعريف وظيفة الهدف.

٥. خوارزميات البحث التطورية يمكن أن توفر عدد من الحلول الممكنة لمشكلة معينة، والإختيار النهائي للمستخدم (Ashlock, 2008).

٣،٣ معاملات الخوارزميات التطورية

٣،٣،١ الإختيار Selection

الإختيار يحدد الأفراد الذين سيتم إختيارهم للتزاوج (التهجين)(Recombination)، وكم عدد أبناء كل فرد مختار ينتج. وتتمثل الخطوة الأولى في اللياقة Fitness عن طريق:

١. إحالة اللياقة التناسبية Proportional Fitness Assignment
٢. رتبة القائم على إحالة اللياقة.
٣. ترتيب متعدد الأهداف.

الإختيار الفعلي يتم في الخطوة التالية، ويتم اختيار الوالدين وفقا للياقتهم عن طريق إحدى الخوارزميات التالية: (Ashlock, 2008)

١. الإختيار بعجلة الروليت Roulette-Wheel Selection
٢. أخذ عينات عشوائية عالمية Stochastic Universal Sampling
٣. الإختيار المحلي Local Selection
٤. إختيار البتر Truncation Selection
٥. إختيار البطولة Tournament Selection .

٣،٣،٢ التهجين Recombination

التهجين ينتج أفراد جديدة تجمع الصفات والمعلومات الواردة من الآباء (الآباء – تزاوج المجتمع)، ويتوقف ذلك على تمثيل المتغيرات من الأفراد، ويمكن تطبيق الخوارزميات التالية:

١. عرض الجميع All Presentation
   ١.١. التهجين المنفصل (المتقطع).
٢. تهجين القيم الحقيقية Real valued Recombination
   ٢.١. التهجين الوسيط.
   ٢.٢. خط التهجين.

المنارة للاستشارات
www.manaraa.com

٢.٣. تمديد خط التهجين.

٣. تهجين القيم الثنائية Binary Valued Recombination

٣.١. نقطة واحدة / مزدوجة / نقاط متعددة متداخلة.

٣.٢. التداخل الموحد.

٣.٣. تداخل خلط ورق اللعب.

٣.٤. التداخل مع البديل المخفض.

تهجين المتغيرات ذات القيم الثنائية يطلق عليه عادة اسم "التداخل" Crossover، ولهذه التسمية أسباب تاريخية؛ فالخوارزميات الجينية تستخدم في الغالب متغيرات ثنائية سميت بالمتداخلة، وكلا المفهومين (التهجين والتداخل) متساويين ويدلان على المعنى نفسه في مجال خوارزمية التطوير (Ashlock, 2008).

٣،٣،٣ الطفرة (التغير) Mutation

وبعد التهجين يخضع كل من الأبناء إلى عملية التغير، تمثيل المتغيرات يحدد الخوارزمية المستخدمة، وأوضح عمليتين هما:

١. عملية التحور لقيم المتغيرات الحقيقية.

٢. عملية التحور لقيم المتغيرات الثنائية (Ashlock, 2008).

٣،٣،٤ إعادة الدمج Reinsertion

بعد إنتاج ذرية يتم إدخالها الى المجتمع. ويتسم هذا بأهمية خاصة إذا نتجت ذرية أقل من الحجم الأصلي للمجتمع. وفي حالة أخرى، ليس كل الأبناء فيها مستخدمين في كل جيل، أو في حال ما إذا كان هناك المزيد من الأبناء المولدة دون الحاجة إليها.

وبموجب خطة إعادة الدمج تم تحديد الأفراد الذين ينبغي إدخالهم إلى المجتمع الجديد، وتحديد الأفراد الموجودة أصلا في المجتمع ويجب أن يحل محلها أفراد جديدة من الأبناء. وخوارزمية الإختيار المستخدمة تحدد مخطط إعادة الدمج:

المنارة للاستشارات
www.manaraa.com

١. إعادة الدمج العالمي لجميع المجتمعات على أساس خوارزمية الإختيار (الإختيار بعجلة الروليت – الإختيار بأخذ العينات العشوائية العالمية – البتر).
٢. إعادة الدمج المحلي لخوارزمية الإختيار المحلية (Ashlock, 2008).

٣،٣،٥ نماذج المجتمع – بالتوازي مع تنفيذ الخوارزميات التطورية.

وتسمح بتعريف إمتدادات الخوارزمية التطورية. هذه الإمتدادات يمكن أن تسهم في زيادة أداء الخوارزميات التطورية. ويمكن التمييز من خلالها بين:

١. النماذج العالمية.
٢. النماذج المحلية (diffusion model, neighborhood model, fine grained mode).
٣. النماذج الإقليمية (migration model, island model, coarse grained model), (Ashlock, 2008).

٣،٣،٦ تطبيق إستراتيجيات متعددة / مختلفة والمنافسة بين المجتمعات الفرعية.

إستنادا إلى نموذج المجتمع الإقليمي، فإن تطبيق إستراتيجيات مختلفة ومتعددة في الوقت نفسه أمر ممكن. ويتم ذلك عن طريق تطبيق مختلف العمليات لكل مجتمع فرعي. ولكفاءة توزيع الموارد أثناء عملية التحسين يستخدم تنافس المجتمعات الفرعية:

١. تطبيق إستراتيجيات متعددة.
٢. المنافسة بين المجتمعات الفرعية.

هذه الإمتدادات من نماذج المجتمعات الإقليمية تساهم في زيادة أداء الخوارزمية التطورية، وخاصة بالنسبة للتطبيقات العالمية الكبيرة والمعقدة (Ashlock, 2008).

المنارة للاستشارات
www.manaraa.com

٣.٤ مقدمة عن الخوارزميات الجينية:

في هذا الفصل سنقدم التقنيات المستخدمة في الخوارزميات الجينية بشكل عام بالإضافة إلى معاملاتها، كيفية عملها وبعض التطبيقات على هذا الحقل من حقول الذكاء الاصطناعي.

٣.٥ تعريف بالخوارزميات الجينية:

الخوارزميات الجينية هي تقنية ذات خاصية مميزة جداً تستخدم في حل المشكلات؛ حيث تقوم بحل المشكلات كما في الطبيعة أكثر من استخدام أسلوب يعتمد على الفرضيات والنظريات (Binti Abdullah,et.al, 2003).

ويوجد العديد من التطبيقات البرمجية التي تعتمد على الخوارزميات الجينية (GA) قد صُممت خصيصاً لإيجاد حلول للمشكلات المستعصية والكبيرة (Belmant, 2001).

وهنالك عدد من المزايا التي تجعل من الخوارزميات الجينية مناسبة لحل المشكلات الصعبة والكبيرة، حيث أن الخوارزميات الجينية صممّت خصيصاً لتقدم أفضل أداء مع المتغيرات غير الخطية ومع البيانات التي تحتوي على أخطاء بالإضافة إلى مقدرتها الكبيرة على البحث في مجموعة كبيرة من الأمثلة المقيّدة بقواعد معينة بحثاً عن الحل الأمثل كما أن لها المقدرة على تحويل القواعد إلى عبارات رقمية ولكن هذه العملية عملية معقدة وتحتاج إلى جهد عالي(Mustafa, 2003)، (Binti Abdullah,et.al, 2003).

المنارة للاستشارات
www.manaraa.com

إن تقيـيم اللياقـة (Fitness) فـي خوارزميـة جينيـة دائمـاً يتطلـب روتـين خـاص بالمشـكلة والنطـاق الـذي سـيتم تطبيـق الخوارزميـة عليـه، لإيجـاد قواعـد خاصـة تمكننـا مـن تحديد المرشحين يجب علينا استخدام بيانات سابقة مناسبة ومنتقاة بشكل جيد.

٣.٦ نقاط رئيسية

- الخوارزميات الجينية (GA) هي محاكاة للتطور الطبيعي.
- الخوارزميـات الجينيـة يمكـن اسـتخدامها لاسـتخلاص أي نظـام سـلوكي والـذي يمكـن إعطـاؤه أي صـفات عمليـة (Operational Characterization) واضحة.
- الخوارزميـات الجينيـة يمكـن أيضـاً ان تسـتخدم فـي حـل المشكلات المبنية على التدريب.

٣.٧ كيفية عمل الخوارزميات الجينية

الخوارزميـات الجينيـة هـي عمليـة حسـابية تحـاكي التطـور الطبيعـي وتحتوي على العمليات الآتية:

١. يـتم بنـاء مجتمـع مـن الأفـراد حيـث يمكـن أن تكـون هـذه الأفـراد إمـا عمـلاء أذكيـاء (Artificial Agents) أو أن تكـون حلـول مقترحـة لمشـكلة محددة.

٢. تقـوم الخوارزميـات الجينيـة بتكـرار عمليـة التـزاوج وإعـادة الإنتاج لإيجاد مجموعة تملك أكبر قدر من اللياقة (Fitness).

٣.٨ الخوارزميات الجينية كـ(pseudo code)

قبـل اسـتخدام الخوارزميـات الجينيـة يجـب أن نجـد نـوع جينـي (Genotype) مناسـب للسـلوك أو لطبيعـة المشـكلة التـي نريـد ان نجـد لهـا حـل وعلينـا أيضـاً أن نجـد

المنارة للاستشارات
www.manaraa.com

معادلة لياقة (Fitness Function) أي طريقة تمكننا من حساب اللياقة لكل جين(كروموسوم) (Binti Abdullah,et.al, 2003)،(Loreana,2001).

الخطوات الرئيسية في الخوارزميات الجينية هي كما يلي:

١. إيجاد مجتمع بدائي من النوع الجيني.

٢. حساب اللياقة لكل فرد في المجتمع.

٣. السماح لعدد محدد من الأفراد ذوي الكفاءة المناسبة بالتزاوج؛ أي إيجاد نسل من هذه الأفراد (Crossover).

٤. إدخال المجموعة الناتجة إلى المجتمع مع إمكانية إزالة عدد مكافئ من الأفراد غير الكفؤة.

٥. تكرار الخطوات من الخطوة الثانية حتى الخطوة الرابعة حتى يتم إيجاد عدد من الافراد التي تملك مقدار مناسب من اللياقة (Fitness) (Loreana,2001).

٣.٩ مشكلات تؤخذ بعين الاعتبار عند تطبيق الخوارزميات الجينية

١. استخدام الخوارزميات الجينية يحتم اختيار شروط البدء والهيكلية المبني عليها بشكل حذر.

٢. يجب اختيار المطابقات رمزية (Encoding Scheme) المستخدمة بشكل دقيق(As: binary scheme, integer scheme,).

٣. يجب اختيار معاملات الدمج والتزاوج.

٤. يجب تجهيز معادلة اللياقة بشكل مناسب للمشكلة التي نقوم بالبحث عن حل لها.

٥. يجب إيجاد طرق مناسبة للحفاظ والحصول على حجم المجتمع المناسب(Loreana,2001).

المنارة للاستشارات
www.manaraa.com

٣.١٠ مشكلات تخص حجم المجتمع

اذا تم التزاوج بإستخدام التداخل من نقطة واحدة (Single Point Crossover) بدون حدوث تحوّل (Mutation) العملية ستفشل إلا إذا كان جميع إجزاء الجين المناسب موجودة في المجتمع البدائي (Loreana,2001).

٣.١١ دالة اللياقة (Fitness Function)

إن نجاح الخوارزمية الجينية مربوط بشكل رئيسي مع دقة دالة اللياقة، فإذا كانت دالة اللياقة غير دقيقة ستكون نتائج الخوارزمية غير دقيقة وبالتالي ستفشل في إيجاد الحل الأمثل ومن هذا المنطلق يجب أن تبنى دالة اللياقة بشكل يناسب فضاء البحث (Search Space) والمشكلة التي نحاول إيجاد الحل لها(Mustafa, 2003)،(Belmant, 2001).

ولذلك نقترح إستخدام دالة لياقة تقوم بإيجاد لياقة الفرد بناءا على المجتمع الذي ينتمي له ومن هنا فإن لياقة المجتمع يمثل مجموع لياقة الأفراد الذين ينتمون إلى هذا المجتمع.

ان دالة اللياقة المستخدمة تعبر عن مجموع المردود للعناصر الممثلة للكروموسوم ويعرف المردود بالمعادلة التالية

المردود=(سعر البيع- سعر الشراء)× الكمية

٣.١٢ معاملات الخوارزميات الجينية

٣.١٢.١ الطفرة Mutation

في معظم الحالات تستخدم الخوارزميات الجينية عمليات الطفرة (تغيير) بالإضافة الى التداخل (Crossover) لإيجاد نسل (Offspring) جديد بشكل تلقائي

المنارة للاستشارات
www.manaraa.com

وهذا قد يساعد في عملية إيجاد جينات رئيسية (تسلسل من الجينات التي لها تأثير مباشر على مستوى اللياقة) والتي ليست موجودة في أي مكان في المجتمع البدائي، وفي العادة هذه العمليات تستخدم الاحتمالات والتبادل، فعلى سبيل المثال:

حينما يكون الترميز المستخدم في بناء الجين أو الكروموسوم هو تمثيل ثنائي فالتغيرات ممكن أن تكون عن طريق قلب الاصفار الى واحدات او العكس وذلك بالاعتماد على الاحتمالات المستخدمة.

الفكرة الرئيسية هنا هي أن الطفرة (Mutation) تقدم تغيرات جينية في المجتمع، فبدون الطفرة (Mutation) يجب على جين الحل أن يكون موجود من البداية في المجتمع البدائي (Belmant, 2001).

٣.١٢.٢ معامل التداخل Crossover

إن عملية التداخل المستخدمة في هذه الدراسة تقوم بشكل اساسي على إعادة الإنتاج عن طريق التداخل من نقطة واحدة (Single Point crossover) وفي مثل هذا السياق يكون النوع الجيني عبارة عن تسلسل من الأحرف عادةً أصفار و واحدات.

إن إتحاد زوج من الجينات عند نقطة القطع يولد طرفان، طرف أيمن وطرف أيسر. والنسل الناتج يكون عبارة عن إتحاد الطرف الأيمن للجين الأول مع الطرف الأيسر للجين الثاني أو العكس أو كلتاهما.

تتم عمليه التداخل بين الكروموسومات بحيث تحدد نقطه فصل بين كروموسومين بشكل تلقائي فنأخذ القسم الأول من الكروموسوم الاول والذي يحتوي على افضل عناصر من الكروموسوم الاول ونقوم بأخذ العدد المتبقي من الكروموسوم الثاني بحيث يكون المجموع الكلي للعناصر في الكروموسوم مساوي للعدد المطلوب.وهذه العملية ممثلة بالشكل (٣-١) تطبيق عملية التداخل Crossover، حيث انه هنالك نسلان C و D والآلية المستخدمة هنا هي Phenotype والوصف الجيني

المنارة للاستشارات
www.manaraa.com

للآلية هو Genotype .والإنتاج النوعي لوالدين يتم عن طريق دمج جيني لهما وهذا بالعادة يتضمن تطبيق عملية التداخل (Crossover) والتي تقوم بربط طرفي النهاية لكل جين (والد) والعكس صحيح. (Loreana,2001).

A(first parent)							
1	1	0	1	0	0	1	1
B(other parent)							
1	0	1	1	0	1	0	1
We select random single crossover point suppose 4							
C (offspring)							
1	1	0	1	0	1	0	1
D (offspring)							
1	0	1	1	0	0	1	1

الشكل(٣-١) تطبيق عملية التداخل Crossover

٣.١٢.٣ **القلب وإعادة التجميع Inversion And Recombination**

القلب Inversion هي عملية قلب جزء من الكروموسوم أما إعادة التجميع Recombination هي عملية تقوم على تغيير إتجاه هذا الجزء ودمجه مع الكروموسوم الناتج. كما في الشكل (٣-٢).

Chromosome

Before inversion : 0010010110101001010110100101010001010

During inversion: 0010010110101001010110100101010001010

1010110100 1

One portion is inverted:

0101001011 0

Recombination: 0010010110101000101001011001010001010

المنارة للاستشارات
www.manaraa.com

After inversion: 0010010110101000101001011001010001010

الشكل(٣-٢) القلب وإعادة التجميع

٣.١٣ تطبيقات الخوارزميات الجينية

هنالك العديد من التطبيقات على الخوارزميات الجينية، ومن هذه التطبيقات ما يلي:

١. الشبكات العصبية (Neural Networks).
٢. جدولة الوقت (Timetabling).
٣. حل المشكلات المثلى (Optimization Problems).
٤. معالجة الصور (Image Processing).
٥. البرمجة الجينية (Genetic Programming).
٦. التنقيب عن البيانات (Data Mining).

٣.١٤ حقول أخرى من تعلم الآلة

٣.١٤.١ تجميع البيانات (Data Clustering)

تجميع البيانات هي عملية وضع البيانات في مجموعات متماثلة. خوارزمية التجميع تقسم مجموعة من البيانات إلى عدة مجموعات. حيث أن التشابه بين النقاط ضمن مجموعة معينة أكبر من التشابه بين نقطتين ضمن مجموعتين مختلفتين. ان تجميع البيانات هي فكره بسيطة في طبيعتها وهي قريبه جدا من الإنسان في طريقة تفكيره حيث اننا كلما تعاملنا مع كميه كبيره من البيانات نميل إلى تلخيص الكم الهائل من البيانات إلى عدد قليل من المجموعات او الفئات،وذلك من اجل تسهيل عمليه

المنارة للاستشارات
www.manaraa.com

التحليل.خوارزميات التجميع تستخدم على نطاق واسع ليس فقط لتنظيم وتصنيف البيانات وانما هي مفيده لضغط البيانات وبناء نموذج ترتيب البيانات. حيث أنه اذا كان بامكاننا أن نجد مجموعات من البيانات، فانه بالامكان بناء نموذج للمشكلة على أساس تلك المجموعات.هناك عدد من التقنيات المستخدمة في عملية تجميع البيانات، ومن هذه التقنيات(الخوارزميات):

- K-Means Clustering
- Fuzzy C-Means Clustering
- Mountain Clustering
- Subtractive Clustering (Mustafa, 2003).

٣.١٤.٢ التحقيق المتقاطع (Cross Validation)

التحقيق المتقاطع(Cross Validation) عبارة عن عمليات حسابية واحصائية لتقسيم البيانات والعينات (samples) الى مجموعات فرعية. ويتم اجراء عملية التحليل في البداية على مجموعة فرعية واحدة، في حين يتم الاحتفاظ بالمجموعات الفرعية الأخرى لاستخدامها لاحقا للتأكد من صحة التحليل الأولي (Mustafa, 2003).

المجموعة الفرعية الأولى يطلق عليها مجموعة التدريب(Training Set)، والمجموعات الفرعية الأخرى يطلق عليها مجموعات الاختبار والتحقق(Validation And Testing Sets) (Mustafa, 2003).

المشكلة في استخدام الطرق السابقة والسالفة الذكر أنها لا تعطي مؤشرا عن كيفية تصرف المعلم (Learner)، عندما يطلب منه التنبؤ حول البيانات التي لم تكن

المنارة للاستشارات
www.manaraa.com

جـــاهزة حتـــى الآن. لا تقـــوم بعمليـــه التعمـــيم لجميـــع البيانـــات المســـتقبلية (Not Generalized) (Mustafa, 2003).

يمكـن التغلـب علـى هـذه المشـكلة مـن خـلال عـدم اسـتخدام مجموعـة البيانـات بالكامـل فـي عمليـه التـدريب، بعـض البيانـات نقـوم بحـذفها قبـل إجـراء عمليـة التـدريب، وبعـد إجـراء عمليـة التـدريب، البيانـات التـي تـم حـذفها يـتم اسـتخدامها فـي عمليـة الاختبـار، وذلـــك لحســـاب الأداء (Performance)، وهـــذه هـــي الفكـــرة الرئيســـية للـــ(Cross Validation) (Mustafa, 2003).

المنارة للاستشارات
www.manaraa.com

٣،١٥التنقيب عن البيانات

هنالك كميات هائلة من البيانات توزع حول العالمِ ومن الصعب جداُ الوصول إلى كل هذه البيانات، والنمو الضخم للبيانات سَيُؤدّي إلى وضع تزيد فيه صعوبة الوصول إلى المعلومات المطلوبة أكثر فأكثر، لكن بدعمِ الحاسبات الحديثة سيمكننا ذلك من أَنْ نجد أيٍ من المعلومات التي نبحث عنها آلياً من أي قاعدة بيانات، هذه العملية - بالطبع - مثيرة جداً للاهتمام، ويطلق على هذه العملية اسم التنقيب عن البيانات.

٣،١٦ ما هو التنقيب عن البيانات

عموماً، التنقيب عن البيانات (وأحياناً تسمى باكتشاف المعرفةَ أَو البيانات) هي عملية تحليل بياناتِ من منظورات مختلفة وتخليصها ومن ثم تحويلها إلى معلوماتِ مفيدةِ (معلومات يمكن أن تُستَعملَ لزيَأْدَة الدخلِ، خفض التكاليف، أَو كلتاهما).

برامج التنقيب عن البيانات تعتبر من الأدوات التحليلية التي تقوم على تحليل البياناتِ، وتسمح للمستعملين بتَحليل البيانات بأكثر من طريقة ومعاير وتمكنه من تصنيفها وتلخيص العلاقات المستخلصة من بينها(Chen, 2000).

وكتعريف لتنقيب عن البيانات هو البحث واكتشاف المعلومات الجديدة من ناحية الأنماط أَو القواعد المأخوذة من كمية كبيرة من البيانات السابقة. ولكي يكون الأمر مفيداً من الناحية العملية،التنقيب عن البيانات يجب أن يُنفّذ بشكل كفوء على الملفات الكبيرة وقواعد البيانات (Chung, 1999).

المنارة للاستشارات
www.manaraa.com

٣،١٧ التنقيب عن البيانات من الناحية التاريخية

التنقيب عن البيانات ظهر في البداية كمنطقة دراسة مستقلة منذ ما يقارب العقد، ولكن حقل تعلم الآلة ظهر من زمن بعيد.ففي الستينات من القرن الماضي وفي بداية السبعينيات من نفس القرن كان حقل تعلم الآلة حقلاً صغيراً وسريع النمو وهو احد حقول الذكاء الاصطناعي.وفي خلال فترة الثمانينات بدأ يحصل فصل بين تعلم الآلة المبني على العلوم النفسية وتعلم الآلة المبني على العلوم التكنولوجية ومن هنا بدأ الأول بالتلاشي مع مرور الوقت (Chen,2000).

٣،١٨ ظهور التنقيب عن البيانات

الثورة السريعة للبيانات المتاحة عبر الشبكة الإلكترونية في أواخر الثمانينات وأوائل التسعينات أعطت فرصة جديدة لتعلم الآلة المبني على العلوم التكنولوجية، وتبين بشكل سريع أن هذا الحقل من الذكاء الاصطناعي سيمكن من استخراج وإيجاد الأنماط والخصائص المشتركة بين البيانات في قواعد البيانات الضخمة وإمكانية استخراج المعلومات منها بشكل أفضل وبقدر أكبر من الكفاءة.وفي هذه الفترة بدأ حقل التنقيب عن البيانات بالظهور كفرع لتعلم الآلة ومن هنا بدأ بالنمو والتطور السريع، وفي أحد التقارير كتقرير (Gartner) المشهور عالمياً تم تقييم التنقيب عن البيانات كأحد أكبر المجالات وأهمها في المستقبل القريب(Siddharth,2002)،(Chung,1999).

٣،١٩ التنقيب في سجلات مركز مواد تموينية

إن استعمال آلات تفقد السلع الأوتوماتيكية في مراكز التموين الحديثة يمكن من حفظ سجلات كبيرة عن عمليات البيع التي تمت في هذه المراكز والتخفيضات وأعداد السلع التي توجد فيها، الخ. تجسد هذه السجلات كميات واسعة مِن البيانات ومن الصعب جداً تحليلها بالطرق التقليدية.

المنارة للاستشارات
www.manaraa.com

أدوات التنقيب عن البيانات المأخوذة من طرق تعلم الآلة يمكن أن تطبق لكي تجد أنماط سلوكية متشابهه بين الزبائن كشرائهم لسلع معينة بكميات معينة ومتى تتم عمليات الشراء تلك والعوامل التي تؤثر على تردد الزبائن إلى المركز التمويني.على سبيل المثال ، أداة تنقيب عن البيانات قد تثبت أن الزيادة في شراء سلعة البوظة قد يؤدي إلى تراجع بيع سلعة الجبنة أو الحليب.

ومعرفة هذا النوع من السلوك يؤهل المركز التمويني أو الشركة أو أي طرف يستخدم أداة تنقيب عن البيانات من أن يزيد أرباحه بشكل ملحوظ وذلك أما بإعادة هيكلة الترتيب والمظهر العام في المركز التمويني أو تطوير السلوكيات المتبعة في الشركة وكل ما يمكن أن نستنتجه من التنقيب عن البيانات واستخدامه لتحسين الإنتاجية من منطلق دراسة هذه البيانات وبناء نتائج واستراتيجيات عليها.

٣،٢٠ عملية التنقيب عن البيانات

التنقيب عن البيانات يتضمن استخراج المعرفة من ضمن عدد كبير من قواعد البيانات التي قد يصل عددها إلى أكثر من مليون أو استخراج المعرفة من أي مصدر معلومات،وفي الحياة العملية، المعرفة تعبر عن حقائق تؤثر على الفوائد بشكل مباشر وقد تزيدها بشكل هائل.

هذه الحقائق يجب أن تسهل عملية التوقع وأن تكون جديدة ولا يمكن استخراجها بسهولة من دون عملية التنقيب عن البيانات، كما أنها يجب أن تكون سهلة الفهم للمستخدم النهائي والذي قد لا يملك أدنى فكرة عن تكنولوجيا المعلومات أو الحقول التي يتضمنها من خوارزميات البحث أو الذكاء الاصطناعي أو التنقيب عن البيانات بحد ذاتها (Chen, 2000).

وعملية اكتشاف المعرفة في قواعد البيانات تأخذ كمدخل مجموعة من قواعد البيانات وتأتي بمخرج يمثل المعرفة العملية.

المنارة للاستشارات
www.manaraa.com

وبشكل عام، فهي تتكون من تطبيق متكرر لثلاث عمليات وهذه العمليات هي

١- التنظيف و إزالة التشويش أو التكرار و إزالة الأخطاء ويتم إزالة البيانات غير المطابقة والناقصة وثم يتم دمج جميع البيانات المستخلصة من مجموعة مصادر غير مترابطة للحصول على مظهر كامل وشامل ومود للحالة التي نريد استخلاص المعرفة لها، والناتج من هذه العملية يدعى بمخزن البيانات Data Warehouse.

٢- اختيار عينة من مخزن البيانات بناءاً على استراتيجيات معطاة، وعملية الاختيار تلك تبنى على أساس أن يكون الناتج يحقق مبتغى عملي يحقق مبتغى الشركة أو المؤسسة المعنية.

٣- تطبيق عملية التنقيب عن البيانات على البيانات المختارة.وفي العادة، يتم تحويل البيانات المختارة إلى أشكال مناسبة قبل تطبيق عملية التنقيب عن البيانات عليها، وتتم عملية التحويل إما بعمليات الجمع أو الربط (Siddhartha, 2002)، (Chung, 1999).

٣،٢١ خطوات التنقيب عن البيانات

الخطوات الرئيسية المتبعة للتنقيب عن البيانات هي كالآتي:

١- اختيار بيانات الهدف من قاعدة البيانات وفقاً لبعض المعايير والقواعد.

٢- يتم التغيير بالبيانات المختارة من قبل المعالج التابع لقاعدة البيانات لينفذ عليها عمليات أولية مثل الترشيح (Filtering).

المنارة للاستشارات
www.manaraa.com

٣- يتم تحويل البيانات المعالجة إلى شكل ما تسمح للمنقب عن البيانات من تلخيص العلاقات بين السجلات المختلفة التابعة لقاعدة البيانات بسهولة.

٤- في الخطوة الأخيرة يقوم منقب البيانات أو برنامج تنقيب البيانات بتحليل العلاقات والأنماط في بيانات العمليات المخزنة بناءاً على طلبات المستخدم.

من الخطوة الأخيرة يمكننا بسهولة الحصول على معلومات مفيدة قد تساعدنا في الوصول إلى الغاية المبتغاة (Chung, 1999).

٣،٢٢ الجهات المهتمة بالتنقيب عن البيانات

التنقيب عن البيانات يستخدم بالعادة من قبل شركات لها تركيز كبير على المستهلك كمنظمات التسويق وشركات الوساطة المالية وغيرها الكثير، والتنقيب عن البيانات يسمح لهذه الشركات بتحديد العلاقات بين العوامل الداخلية كالسعر أو مهارات طاقم الشركة و العوامل الخارجية كالمنافسة الناتجة من المؤشرات المالية وخصائص الزبائن، فهي تسمح لهم بتحديد التأثير على المبيعات، مثل رضا الزبون وأرباح الشركة وأخيراً فهي تسمح لهم بتحويل البيانات إلى معلومات ملخصة لعرض بيانات مفصلة للعمليات التي تتم في الشركة.

وباستخدام التنقيب عن البيانات يستطيع البائع أن يتتبع سجلات مشتريات زبون وبذلك يستطيع أن يبعث له إعلانات محددة بناءً على تاريخ مشتريات هذا الزبون ورغباته ومن هنا يمكن للبائع أن يحدد ما هي المنتجات التي تعني وتؤثر في الزبون وبذلك يدرك كيفية جذب الزبون وإقناعه بالشراء،فعلى سبيل المثال مركز ترفيهي قد

المنارة للاستشارات
www.manaraa.com

ينقب في بيانات تأجير الأفلام لديه ليستطيع أن يحدد لزبائن ويوصيهم بمشاهدة أفلام محددة بناءً على أختياراتهم السابقة، وموزع بطاقات ائتمانية قد يقترح لزبائنه شراء منتجات محددة بناءً على مقدار المصاريف الشهرية لهم.

٣،٢٣ كيف يعمل التنقيب عن البيانات؟

بينما تتطور تكنولوجيا المعلومات فإن التنقيب عن البيانات في أنظمة التحليل كانت تؤمن الوصل بين نظام تحليل علاقات التنقيب عن البيانات والأنماط الموجودة في العمليات المخزنة المبنية على طلبات المستخدم (Murray, 2000)،(Melab, 2001).

والعديد من المحللين بشكل عام أرادوا الاستجابة مع أربع أنواع من العلاقات وهي:

١-البيانات التقليدية المخزنة تستخدم لتحديد مواقع البيانات في مجموعات محددة بشكل سابق.على سبيل المثال، سلسلة مطاعم قد تنقب بيانات شراء الزبون لمعرفة متى يزورها الزبون وماذا يطلب بالعادة وهذه المعلومات قد تؤدي إلى زيادة الطلب وذلك عن طريق تحديد عروض يومية.

٢-عناصر البيانات المترابطة أو العنقودية يتم تجميعها بناءً على علاقات منطقية أو تفصيلات الزبون،على سبيل المثال، يمكن أن يتم التنقيب عن البيانات لتحديد تقسيمات السوق أو رغبات الزبون.

٣-البيانات المرتبطة يمكن ان يتم التنقيب عنها لإيجاد روابط مع بيانات أخرى.

٤-البيانات ذات الأنماط المدرجة يتم التنقيب عنها من أجل توقع أنماط واتجاهات سلوكية، فعلى سبيل المثال، فبائع معدات خارجية قد يتوقع أحتمال شراء حقيبة ظهر بناءً على شراء زبون أحذية السفر وأكياس نوم (Melab, 2001).

المنارة للاستشارات
www.manaraa.com

٣،٢٤ العناصر الرئيسية في التنقيب عن البيانات

التنقيب عن البيانات يحتوي على خمس عناصر رئيسية وهي:

١- استخلاص بيانات من عمليات التحويل والتحميل ووضعها في نظام تخزين بيانات.

٢- تخزين وإدارة البيانات في نظام يحتوي على أكثر من بعد.

٣- السماح بالوصول إلى البيانات لمحللي الأعمال و مختصو تكنولوجيا المعلومات.

٤- تحليل البيانات من قبل نظام محوسب.

٥- عرض البيانات في هيئة مفيدة كرسم بياني أو جدول(Melab, 2001).

٣،٢٥ اكتشاف المعرفة في قواعد البيانات

ان تطبيق تقنيات التنقيب عن البيانات بصورة عمياء على بيانات خامة هو غير مجدي، فليس فقط يجب أن تجهز البيانات بل يجب تحديد الخوارزمية والتقنية الصحيحة بناءً على الهدف العام من تنفيذ عملية التنقيب عن البيانات بما في ذلك التحضير مقدماً والتحليل لكل النتائج بعدها تسمى باكتشاف المعرفة في قواعد البيانات KDD (Knowledge Discovery in Data bases) (Murray,2000).

المنارة للاستشارات
www.manaraa.com

٣،٢٥،١ الخطوة المعنية باكتشاف المعرفة في قواعد البيانات بشكل عام:

١- الاختيار Selection: وهذه الخطوة تتضمن تحديد السجلات التي يجب أن تزود إلى خوارزمية التنقيب عن البيانات، فكل السجلات يمكن أن تختار من أجل القيام ببحث هائل عن جميع العلاقات والقواعد غير المعروفة وإزالة البيانات غير المعنية وهي في العادة بيانات تقع خارج النطاق المتوقع من البيانات.

٢- المعالجة المسبقة Preprocessing: وهنا نحدد أستراتيجية أما لتقريب حقول البيانات في قاعد البيانات أو إهمال هذه الحقول.

٣- التنقيب عن البيانات Data Mining: وفي هذه النقطة نحن جاهزون لأخذ البيانات وتحليلها بناءً الهدف المرغوب ويتم اختيار خوارزمية لتناسب البيانات المعنية وتستطيع تقديم نوع الناتج الذي نريده ونبحث عنه.

٤- التفسير والتقييم Interpretation and Evaluation: أخيراً يجب القيام ببعض التفسير والتقييم للبيانات الناتجة، فعلى أحد ما أو برنامج ما ان يقوم ببعض الفحوصات الحسابية لمعرفة ما إذا كانت الأنماط المكتشفة ذات علاقة وإزالة الاستنتاجات المتكررة وفي النهاية تقديم نتائج مفهومة للمستخدم النهائي (Murray, 2000)

٣،٢٦ تطبيقات على التنقيب عن البيانات

وسائل التنقيب في البيانات تُستعمل وبنجاح في الكثير من التطبيقات الحقيقة حول العالم. التطبيقات التالية تشمل بعضا من الأمثلة: (يتبعها مثال لكل تطبيق)

- كتابة تقرير مختصر عن فئة معينة Profiling Populations: تطوير وإنشاء تقارير موجزة عن الزبائن المهمين وعن بطاقات الائتمان.

- تحليل النزعة التجارية Analysis of Business Trend: إيجاد الأسواق ذات قدرات النمو القوية أو الضعيفة.

المنارة للاستشارات
www.manaraa.com

- التسويق لفئة معينة Target Marketing: إيجاد الزبائن من أجل منح التخفيضات لهم لسبب معين.

- تحليل الاستعمال Usage Analysis: إيجاد منوال معين لاستعمال الخدمات والسلع

- فعالية الحملة Campaign Effectiveness: مقارنة استراتيجيات الحملات مع بعضها البعض من أجل إيجاد أكثرها فعالية وتأثيرا.

- جاذبية السلعة: إيجاد السلع التي تباع مع بعضها البعض (Murray,2000).

تطبيقات التنقيب في البيانات بدأت تنمو بصورة كبيرة للأسباب التالية:

١) كمية البيانات الموجودة في مخزن البيانات وسوق البيانات تنمو بصورة أسية (Exponential). ومن أجل ذلك، فإن المستخدم يحتاج إلى أدوات متطورة من مثل التنقيب في البيانات من اجل استخلاص الفائدة والمعرفة من هذه البيانات.

٢) الكثير من أدوات التنقيب عن البيانات بدأت تظهر مؤخرا، وكل أداة أفضل من الأخرى.

٣) المنافسة الشديدة الموجودة في السوق تدفع الشركات إلى الاستفادة القصوى من البيانات التي بيدها عمليات التنقيب في البيانات تفعل ذلك تماما.

المنارة للاستشارات
www.manaraa.com

الفصل الرابع

النظام المقترح

المنارة للاستشارات
www.manaraa.com

٤.١ مقدمة

في هذا الفصل سوف نقدم مثال حي على استخدام الخوارزميات الجينية في التنقيب عن البيانات وكيف أن استخدام هذه التقنية يسهل التحليل والإستخلاص للبيانات وتحويلها الى معلومات مفيدة أي معلومات تفيد في زيادة الدخل والفائدة وتخفيض التكلفة.

الفكرة ان لدينا مركز تمويني يحتوي على العديد من المنتجات والسلع ونرغب بوسيلة فعالة ومحوسبة لتطوير أداء وفعالية هذا المركز التمويني من حيث المحتويات والمبيعات وعدد الزبائن وذلك بتحليل البيانات السابقة لهذا المركز التمويني.

٤.٢ الطريقة

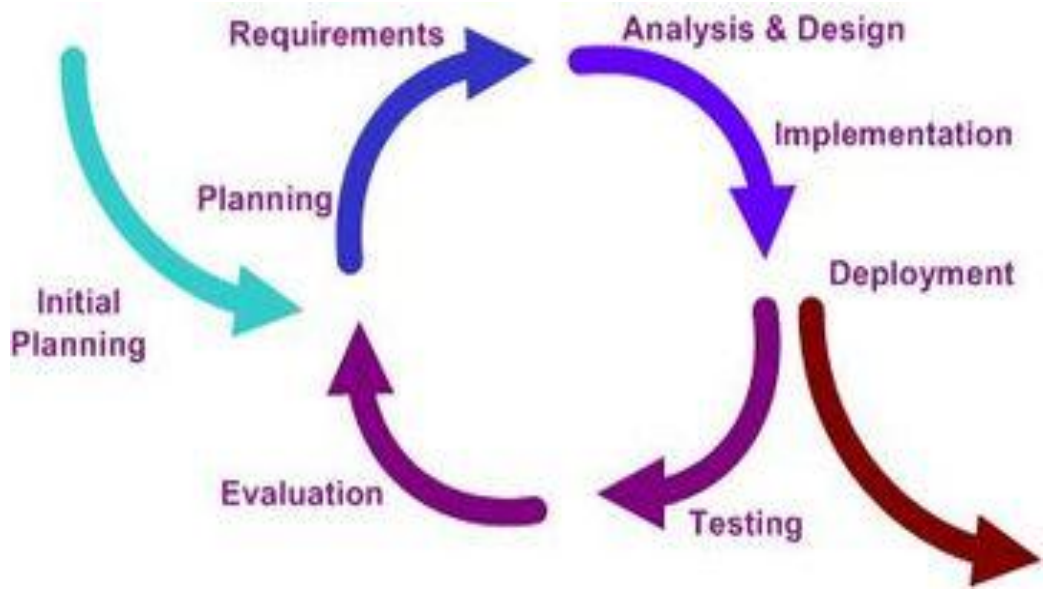
٤,٢,١ طريقة البحث المعتمدة

تم اعتماد طريقة التطوير التكراري والتزايدي (Iterative and Incremental Development) في بناء النظام المقترح، لانها تعتمد تطوير الأنظمة بشكل دوري وتلبية للاحتياجات التي تظهر اثناء عملية بناء النظام وتطويره، وتمكن المبرمج من بناء النظام من النموذج الأولي، وتطوير أجزاء محددة من النظام خلال مراحل العمل المختلفة، وتعطيه فرصة للتعلم من خلال تطبيق واختبار النظام المقترح. ولما هذه الطريقة من مميزات حيث تعطي للمبرمج البدء بعملية بناء النظام دون ان يكون لدية تصورا كاملا ونهائيا عن النظام، بحيث يتمكن من التطوير والإضافة على

المنارة للاستشارات
www.manaraa.com

النظـام مـن خـلال البـدء بعمليـة تطبيـق النظـام فـي أي مرحلـة تـم التوصـل اليهـا واجـراء اختبار له ، وبناء على عملية الاختبار يتم اضافة وظائف جديدة (Alistair, 2008).

ويـبن الشـكل (٤-١) مراحـل التطـوير التكـراري والتزايـدي، ومـن خـلال الشـكل تظهـر المراحـل الأساسـية حيـث تبـدأ مرحلـة التطـوير مـن مرحلـة التخطـيط الأولـي(Initial Planning) ثـم مرحلـة التحليـل والتـــصميم ويتبعهـا الاختبـار ثـم تقيـيم النظـام ثـم العـودة الـى الخطـوة الأولـى ويـتم التكـرار حتـى يـتم الوصـول الـى النمـوذج النهـائي للنظـــام (Alistair, 2008).



الشكل (٤-١) مراحل التطوير التكراري والتزايدي

المنارة للاستشارات
www.manaraa.com

ان الهدف من هذا النظام وهو ان نستطيع ان نتخلص من أي سلعة في المركز التمويني ليس عليها إقبال اوتسبب خسائر او لا تحقق ارباح بالشكل المطلوب لهذا المركز التمويني وتبديلها بسلعة أخرى أفضل منها وتحقق ارباح أفضل للمركز التمويني، لقد أستخدمنا جدول يحتوي على معلومات عن السلع وهذه المعلومات هي كالآتي:

١- رقم السلعة ID

٢- اسم السلعة Item

٣- سعر الشراء Purchase Price

٤- سعر البيع Sell Price

٥- العدد المباع Frequency

٦- الشهر المباع فيه Month

المنارة للاستشارات
www.manaraa.com

وفـي الشـكل التـالي، نسـتطيع ان نـرى مجموعـه مـن المعلومـات الموجـودة فـي قاعدة البيانات :

Id	Item	Selling_Price	Purchase_Price	Frequency	Month
1	حلب	3	1	20	1
2	قهوة	2	1	100	1
3	شوكولاتة ...	5	3	30	1
4	بوظة	3	1	15	1
5	سكر	4	2	150	1
6	شاي	3	1	65	1
7	كتشب ...	7	5	12	2
8	بطاط	6	3	14	2
9	بندورة	4	2	40	2
10	سيريلاك ...	8	6	30	2
11	حليب	3	1	43	2
12	أرز	2	1	20	2
13	قهوة	2	1	10	3
14	شوكولاتة ...	3	1	80	3
15	بوظة	2	1	15	3
16	سكر	3	1	20	3
17	شاي	3	2	15	3
18	كتشب ...	3	1	6	3
19	بطاط	2	1	70	4
20	بندورة	4	1	23	4
21	سيريلاك ...	3	1	33	4
22	حليب	4	2	75	4
23	بوظة	4	2	88	4
24	كتشب ...	4	3	55	4

الشكل(٤-٢) المعلومات المجدولة لسلع المركز التمويني

المنارة للاستشارات
www.manaraa.com

كمـا نعلـم ان هنالـك عـدد مـن المحـددات التـي تواجـه اي مركـز تمـويني عنـدما تكون فكرة تطوير جودة بضائعه، ومن هذه المحددات:

١- ما هي المعايير التي على اساسها يتم ازالة السلع؟

٢- ما هي السلع التي يجب أن تتوفر منها كميات كبيرة؟

٣- ما هو مقدار الزيادة الذي سيؤخذ بعين الإعتبار؟

٤- ما هي السلع التي تشهد اقبالا من قبل الزبائن؟

وغيرهـا الكثيـر مـن المحـددات التـي تواجـه المركـز التمـويني، وفـي هـذا البحـث سـنركز علـى إيجـاد حلـول لمعظـم هـذه المشـاكل عـن طريـق تطبيـق الخوارزميـات الجينيـة في التنقيب عن البيانات.

٤.٣ تحليل النظام المقترح

هنـا سـوف ننـاقش تتطبيـق الخوارزميـات الجينيـة لزيـادة المـردود والـدخل والفائـدة العائدة على المركز التمويني ولحل هذه المشكلة إستخدمنا:

١- Microsoft Visual Studio .NET 2005 كبيئة عمل مستخدمة لبناء النظام

٢- C# كلغة برمجة وبناء الخوارزميات الجينية و VB.NET كلغة برمجه الواجهة الجرافيكيه.

٣- Microsoft SQL Server express 2005 لبناء قاعدة البيانات

وأيضـاً فـي هـذا الجـزء مـن البحـث نقـدم دراسـة لكيفيـة معالجـة المشـكلة بشـكل مهنـي عـن طريـق تحليـل المتطلبـات وبنـاء تصـميم لحـل المشـكلة ومـن ثـم تطبيقـه عملياً.

المنارة للاستشارات
www.manaraa.com

٤.٣.١ دراسة المتطلبات

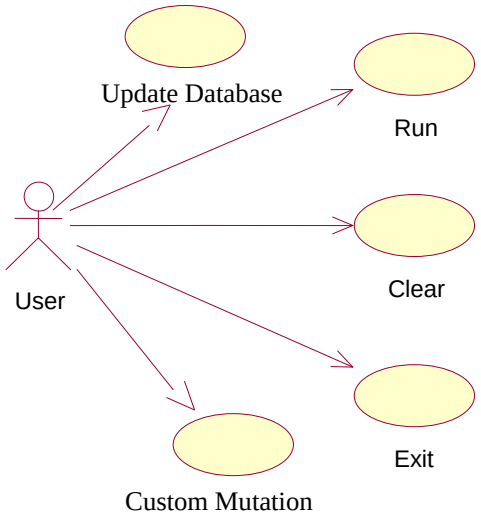
ويتم فيها تحليل المتطلبات وبناء نموذج للنظام المقترح باستخدام لغة النمذجة الموحدة Unified Modeling Language (UML)،وهي لغة سهلة الاستخدام وغنية برسوم قياسية متعددة، لتعطي وجهات نظر متعددة. وتيسر بناء وتطوير النظم.

٤.٣.٢ المستخدمون للنظام

وهو مسؤول المركز التمويني(User) والذي يتفاعل مع النظام.

٤،٣،٢،١ نموذج حالة الاستخدام Use Case Model

وهو مخطط وصف لسلوك النظام من وجهة نظر المستخدم، وهو ذو فائدة خلال مراحل التحليل والتطوير وتساعد في فهم المتطلبات وهو سهل الاستيعاب.



الشكل(٤-٣) نموذج حالة الاستخدام Use Case Model

المنارة للاستشارات
www.manaraa.com

٤.٣.٣ التعريف بالمتطلبات العملية Functional Requirement

١-التشغيل Run:

وتقوم هذه العملية بالربط مع قاعدة البيانات وتحليلها بإستخدام الخوارزميات الجينية الاظهار نتائج تتمثل بأفضل كروموسوم لكل جيل ورسم بياني يمثل مقارنه بين كل جيل من حيث المجموع الكلي لل Fitness .

٢- مسح Clear:

وتقوم هذه العملية بإزالة جميع البيانات السابقة لتمكننا من تتبع كل جيل على حدة.

٣-خروج Exit:

وتقوم هذه العملية بإنهاء البرنامج.

٤- تزاوج اختياري Custom Mutation:

وتقوم هذه العملية بمزاوجة شهرين يحددهما المستخدم .

٥- تجديد قاعدة البيانات Update Database:

تمكن هذه العملية من إضافة سلعه جديدة الى قاعده البيانات.

٤.٣.٤ مواصفات العمليات Specification

يمثل الجدول التالي المواصفات العملية لعملية تشغيل النظام(عملية التنقيب عن البيانات)

المنارة للاستشارات
www.manaraa.com

جدول(٤-١) يمثل عملية تشغيل التنقيب عن البيانات

تشغيل Run
المعنيون: المستخدم User
الوصف: **تقوم هذه العملية بالربط مع قاعدة البيانات والبدء بالتحليل لتحصل لنا على أفضل ترتيب للبضائع.**
التسلسل الطبيعي Basic flow **١- يضغط المستخدم زر Run بعد تحديد القيمة المراد الوصول إليها Optimal** **يقوم النظام بالربط مع قاعدة البيانات وإنشاء الخوارزمية الجينية ويبدأ بالتحليل.**
التسلسل الثانوي Alternative flow **١- في حال عدم إمكانية الربط مع قاعدة البيانات يظهر النظام إشعار للمستخدم.** **٢- في حال عدم تعبئه المستخدم لل optimal يقوم النظام بتحذير المستخدم.**
الشروط القبلية Pre conditions **على المستخدم تشغيل النظام من ملف التشغيل الرئيسي**
الشروط اللاحقة Post conditions **على النظام أن يبدأ بتحليل قاعدة البيانات والقيام بالتنقيب عن البيانات في هذه القاعدة وإظهار النتائج.**

المنارة للاستشارات
www.manaraa.com

يمثـل الجـدول التـالي المواصـفات العمليـة لعمليـة المسـح(عمليـة مسـح النتـائج السـابقة للتنفيذ النظام)

جدول (٤-٢) يمثل عملية مسح المعلومات السابقة

مسح Clear
المعنيون: **مسؤول المركز التمويني** User
الوصف: **تقوم هذه العملية بمسح جميع البيانات السابقة للجيل الحالي.**
التسلسل الطبيعي Basic flow **١- يضغط المستخدم زر** Clear **٢- يقوم النظام بمسح جميع معلومات الأجيال السابقة.**
التسلسل الثانوي Alternative flow **لا يوجد**
الشروط القبلية Pre conditions **على المستخدم تشغيل عملية التنقيب عن البيانات**
الشروط اللاحقة Post conditions **على النظام أن يمسح جميع بيانات الأجيال السابقة**

يمثل الجدول التالي المواصفات العملية لعملية اغلاق النظام

المنارة للاستشارات
www.manaraa.com

جدول (٤-٣) يمثل عملية إغلاق النظام

إغلاق Exit
المعنيون: **مسؤول المركز التمويني** User
الوصف: **تقوم هذه العملية بإنهاء البرنامج.**
التسلسل الطبيعي Basic flow **١- يضغط المستخدم زر** Exit **٢- يقوم النظام بإنهاء عملياته والخرزج من نطاق عمليات نظام التشغيل.**
التسلسل الثانوي Alternative flow **لا يوجد**
الشروط القبلية Pre conditions **على المستخدم أن يكون قد أطلق النظام.**
الشروط اللاحقة Post conditions **على النظام ينهي عملياته ويخرج من نطاق عمليات نظام التشغيل.**

يمثل الجدول التالي المواصفات العملية لعملية التزاوج الاختياري

جدول (٤-٤) يمثل عملية التزاوج اختياري

المنارة للاستشارات
www.manaraa.com

تزاوج اختياري Custom Mutation
المعنيون: **مسؤول المركز التمويني** User
الوصف: **تقوم هذه العملية بمزاوجة شهرين يقوم باختيارهما المستخدم وإظهار نتائج التزاوج.**
التسلسل الطبيعي Basic flow **١- يختار المستخدم الشهرين اللذان يريد إنشاء كر وموسوم من خلالهما** **٢- يضغط المستخدم على رز Custom Mutation.**
التسلسل الثانوي Alternative flow **لا يوجد**
الشروط القبلية Pre conditions **على المستخدم أن يكون قد أطلق النظام.**
الشروط اللاحقة Post conditions **على النظام أن يقوم بمزاوجة هذين الشهرين لإظهار كر وموسوم جديد .**

يمثل الجدول التالي المواصفات العملية لعملية تحديث قاعدة البيانات

جدول (٤-٥) يمثل عملية تحديث قاعدة البيانات

المنارة للاستشارات
www.manaraa.com

تجديد قاعدة البيانات Update Database
المعنيون: مسؤول المركز التموينى User
الوصف: تمكن هذه العملية من إضافة سلعه جديدة الى قاعده البيانات.
التسلسل الطبيعي Basic flow ٣- يضغط المستخدم زر Update Database ٤- يقوم النظام بإظهار شاشه جديدة تمكن المستخدم من اضافه السلعه . ٥- يقوم المستخدم بتعبئة الحقول المعنية و الضغط على زر update. ٦- يقوم النظام بالاضافه على قاعدة البيانات.
التسلسل الثانوي Alternative flow معلومات مضافة خاطئة: يقوم النظام بتحذير المستخدم.
الشروط القبلية Pre conditions على المستخدم أن يكون قد أطلق النظام.
الشروط اللاحقة Post conditions على النظام أن يقوم باضا فه السلعة إلى قاعدة البيانات .

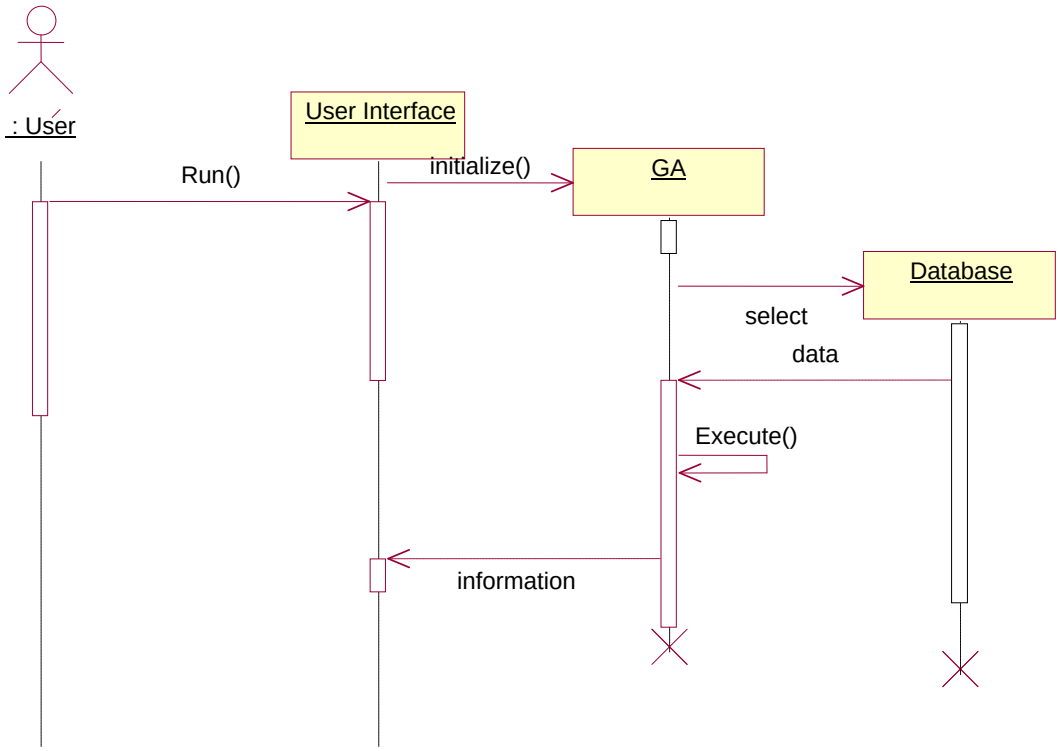
المنارة للاستشارات
www.manaraa.com

٤.٣.٥تصميم النظام

٤،٣،٥،١ مخططات تسلسل الأحداث والتعاون Sequence and Collaboration Diagrams

ومخطط التتابع يبين تسلسل الأحداث لكل عملية حسب التسلسل الزمني لحدوثها في النظام، ومخطط التعاون يعرض التفاعلات التي تحدث للكائنات والوصلات التي بينها وتستخدم الأرقام لبيان تسلسل التفاعلات.

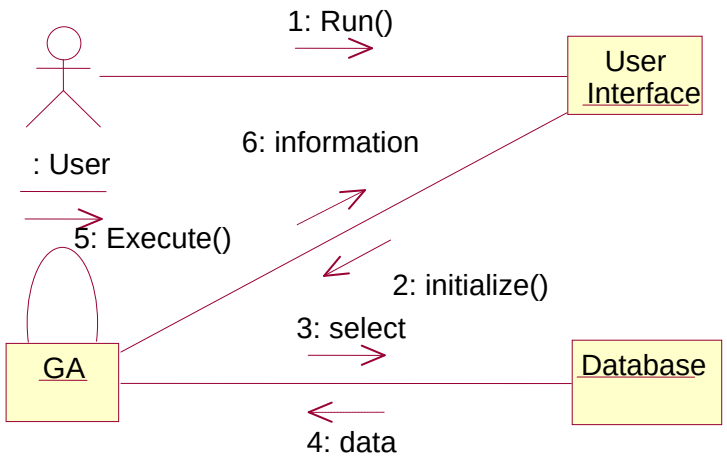
والشكل (٤-٤) يبين تسلسل الأحداث لعملية التشغيل RUN حسب التسلسل الزمني لحدوثها في النظام.



الشكل(٤-٤) تسلسل الأحداث لعملية التشغيل RUN

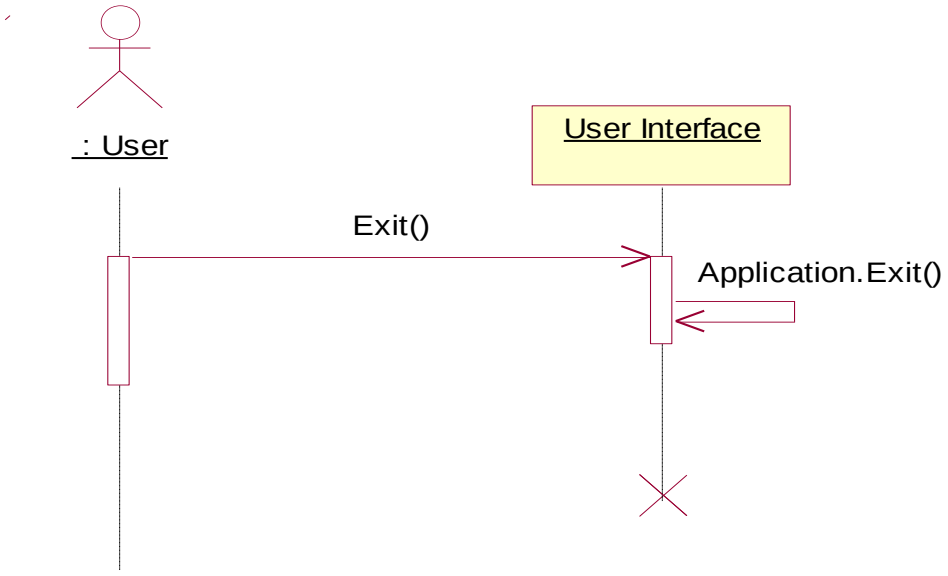
المنارة للاستشارات
www.manaraa.com

والشكل(٤-٥) يبين وصف لكيفية تفاعل الكائنات في عملية التشغيل مع بعضها البعض



الشكل(٤-٥) التعاون في العمليات لعملية التشغيل RUN

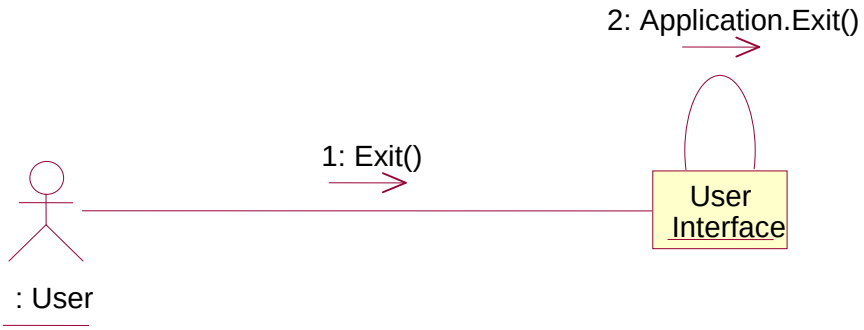
والشكل (٤-٦) يبين تسلسل الأحداث لعملية الإغلاق Exit حسب التسلسل الزمني لحدوثها في النظام



الشكل(٤-٦) تسلسل الأحداث لعملية الإغلاق Exit

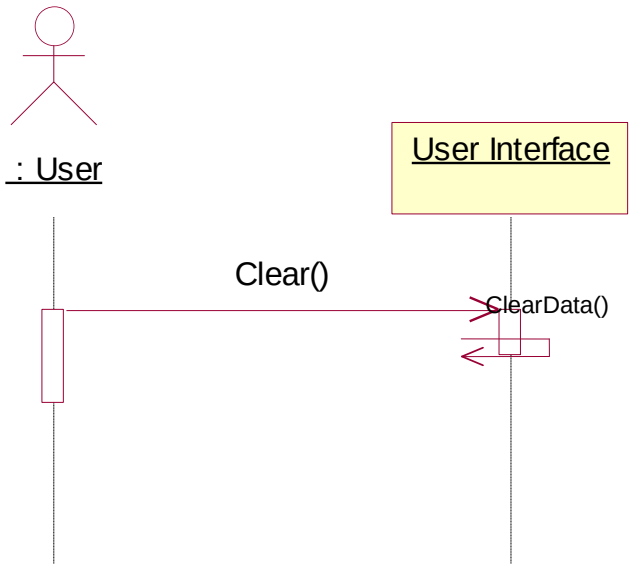
المنارة للاستشارات
www.manaraa.com

والشكل(٤-٧) يبين وصف لكيفية تفاعل الكائنات في عملية الاغلاق مع بعضها البعض



الشكل(٤-٧) التعاون في العمليات لعملية الإغلاق Exit

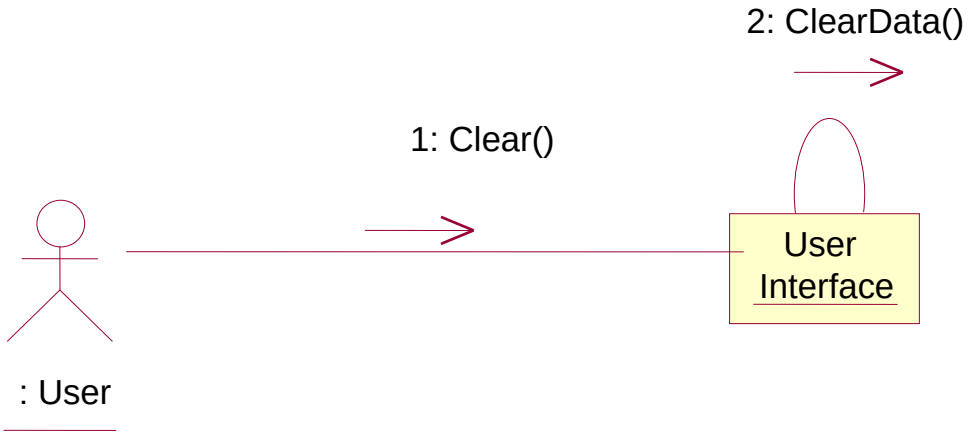
والشكل (٤-٨) يبين تسلسل الأحداث لعملية المسح Clear حسب التسلسل الزمني لحدوثها في النظام



الشكل(٤-٨) تسلسل الأحداث لعملية المسحClear

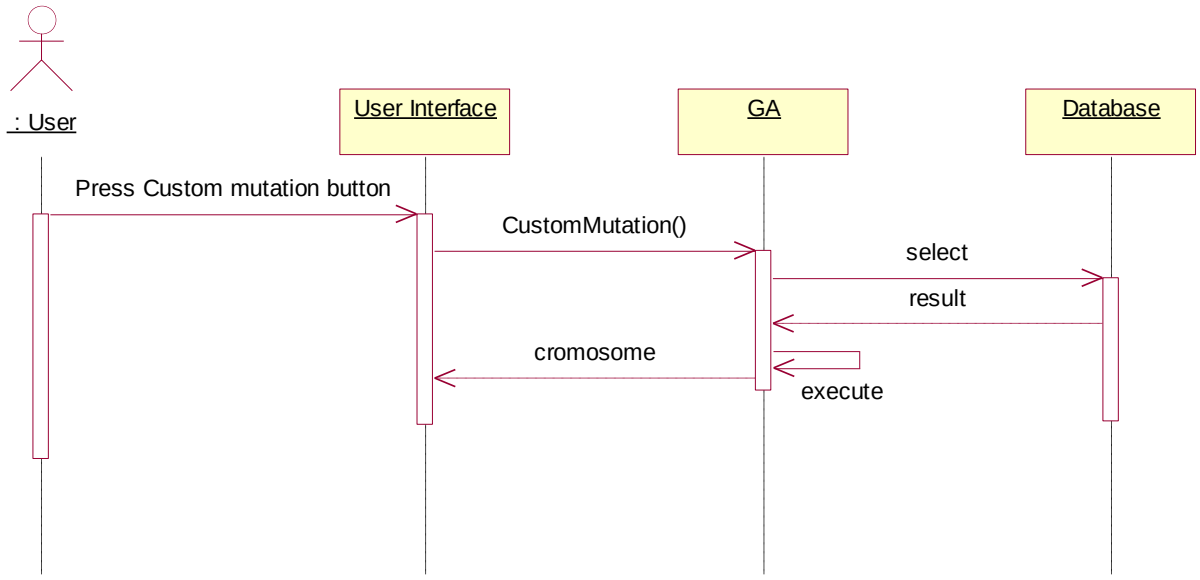
المنارة للاستشارات
www.manaraa.com

والشكل(٤-٩) يبين وصف لكيفية تفاعل الكائنات في عملية المسح مع بعضها البعض



الشكل(٤-٩) التعاون في العمليات لعملية المسحClear

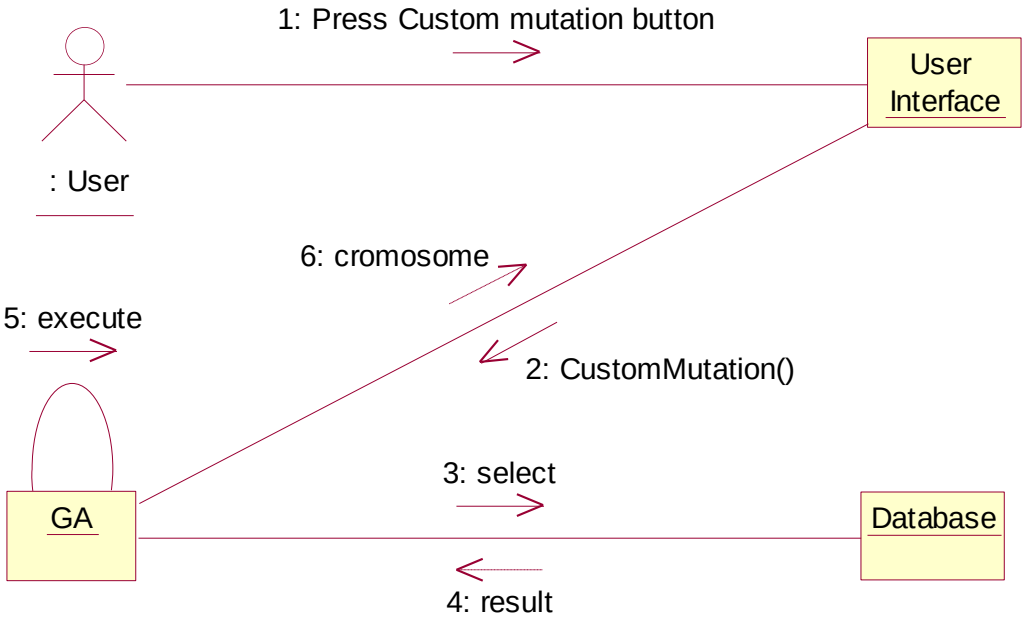
والشكل (٤-١٠) يبين تسلسل الأحداث لعملية التزاوج الاختياري Custom Mutation حسب التسلسل الزمني لحدوثها في النظام



الشكل(٤-١٠) تسلسل الأحداث لعملية التزاوج الأختياري Custom Mutation

المنارة للاستشارات
www.manaraa.com

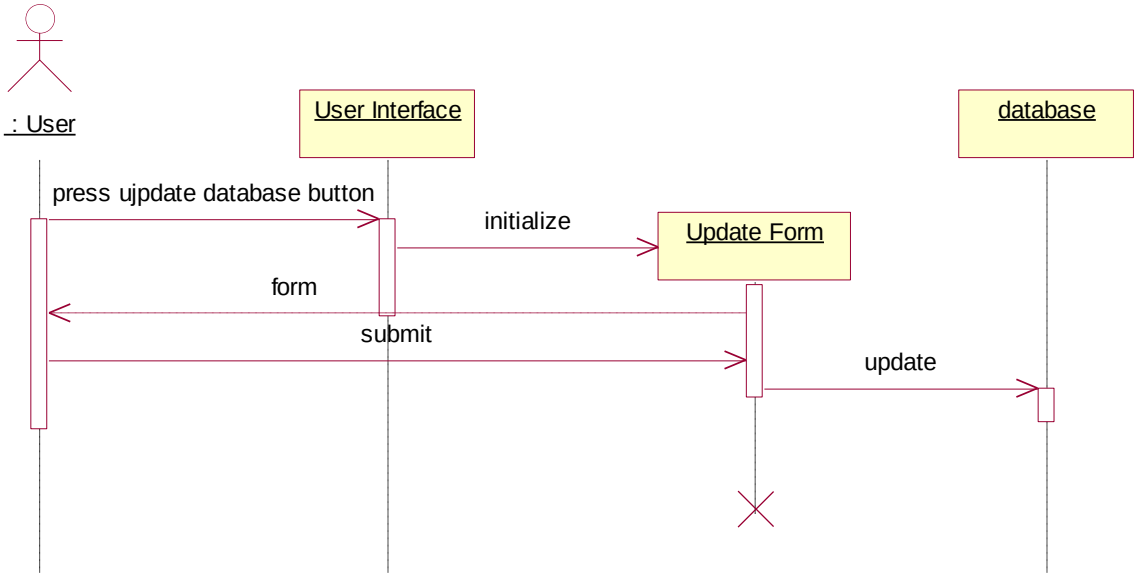
والشكل(٤-١١) يبين وصف لكيفية تفاعل الكائنات في عملية التزاوج الاختياري مع بعضها البعض



الشكل(٤-١١) التعاون في العمليات لعملية التزاوج الاختياريCustom Mutation

والشكل (٤-١٢) يبين تسلسل الأحداث لعملية تحديث قاعدة البيانات Update Database حسب التسلسل الزمني لحدوثها في النظام

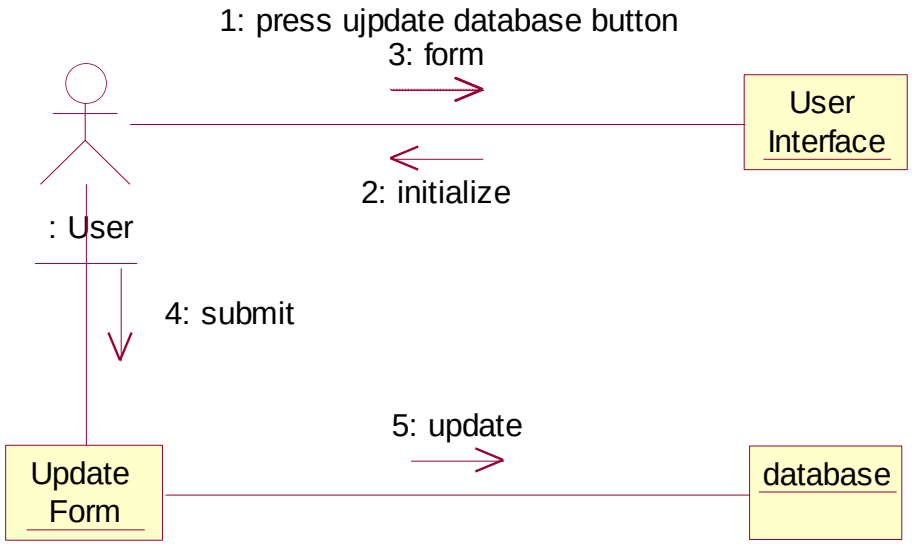
المنارة للاستشارات
www.manaraa.com



الشكل(٤-١٢) تسلسل الأحداث لعملية تجديد قاعدة البيانات Update database

والشكل(٤-١٣) يبين وصف لكيفية تفاعل الكائنات في عملية تجديد قاعدة البيانات مع بعضها البعض

www.manaraa.com

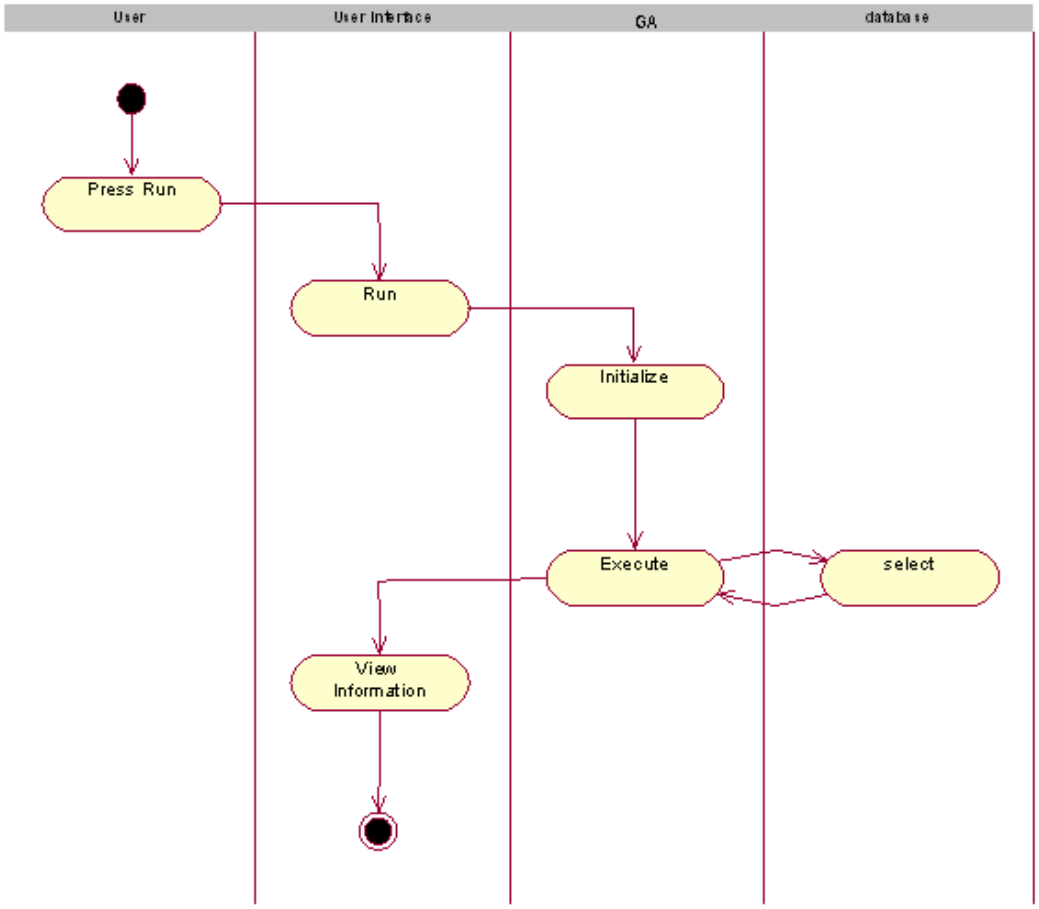


الشكل(٤-١٣) التعاون في العمليات لعملية تجديد قاعدة البيانات Update database

www.manaraa.com

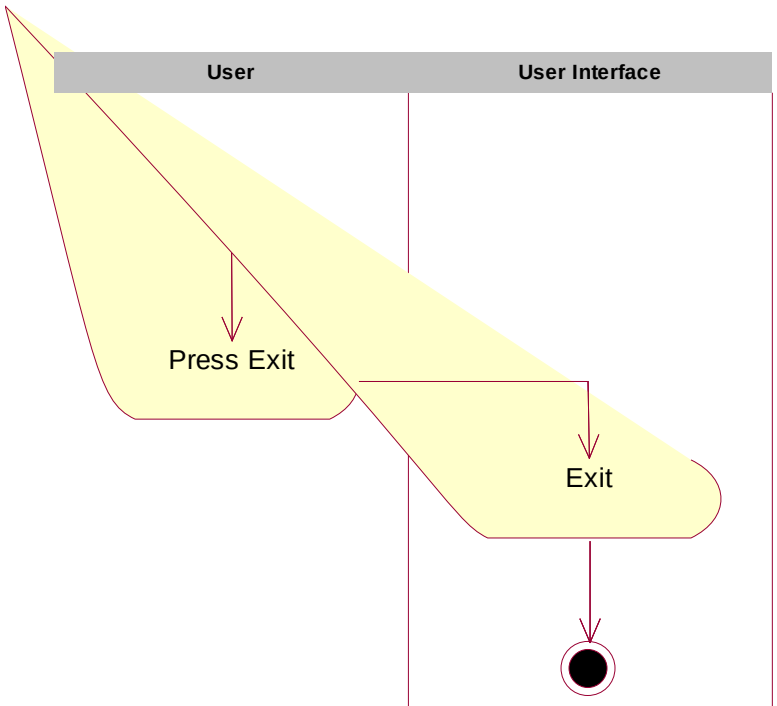
٤.٣،٥،٢ مخططات الأحداث Activity Diagrams

وتمثـل هـذه المخططـات الأحـداث وتوزيعهـا حسـب مكـان حـدوثها وتسلسـلها فـي النظـام.

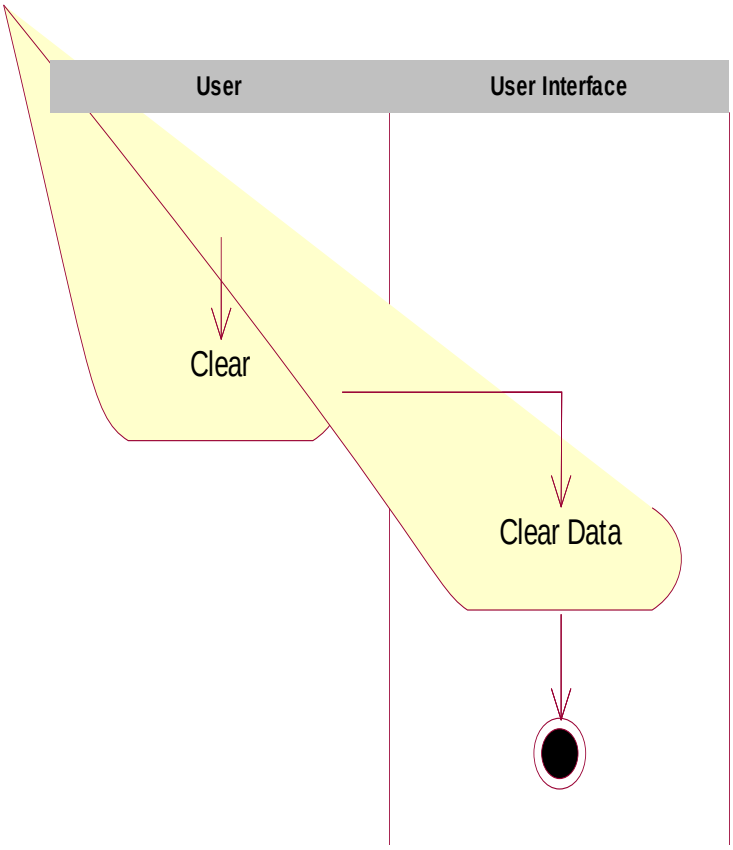


الشكل(٤-١٤) الأحداث لعملية التشغيل Run

المنارة للاستشارات
www.manaraa.com

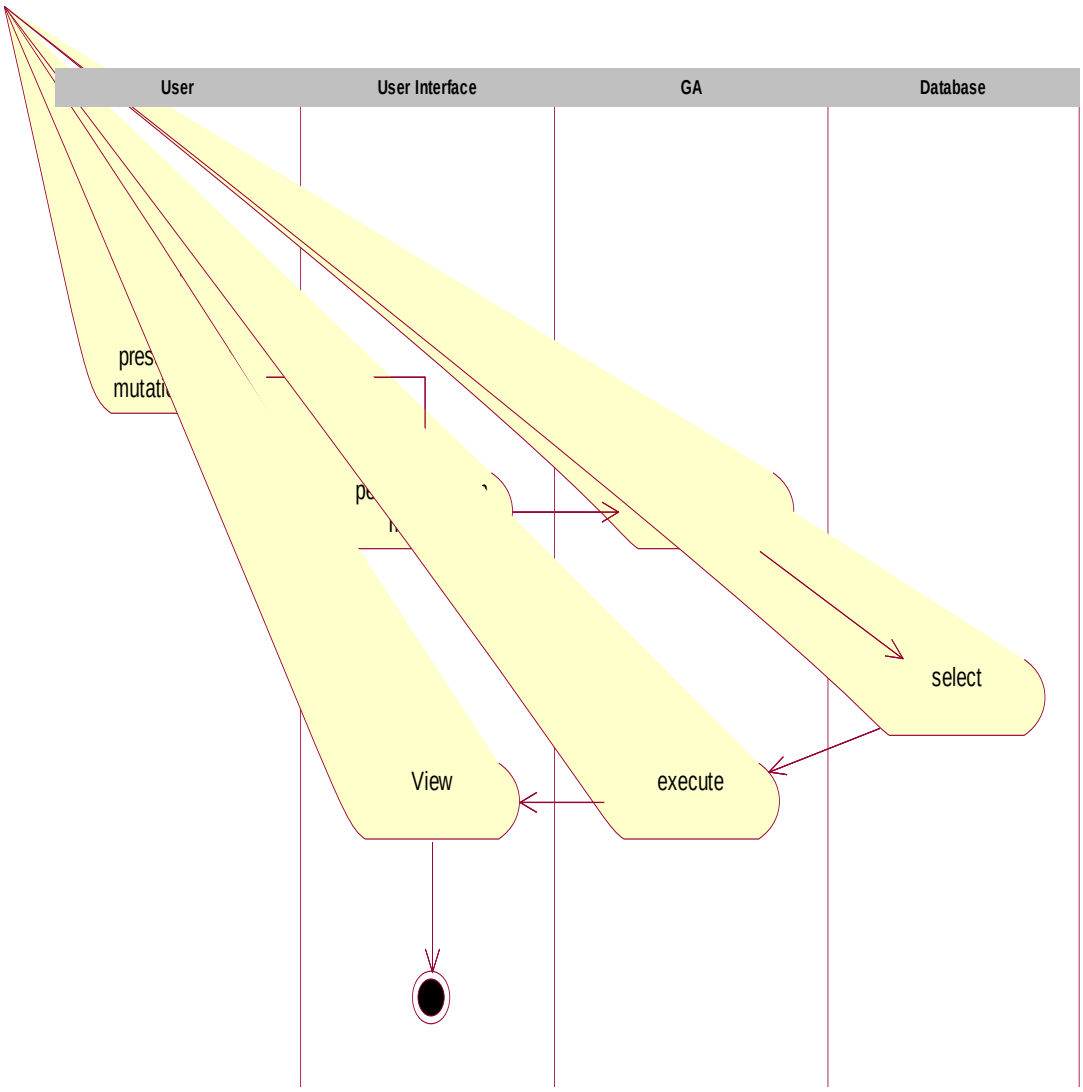


الشكل(٤-١٥) الأحداث لعملية الإغلاق Exit



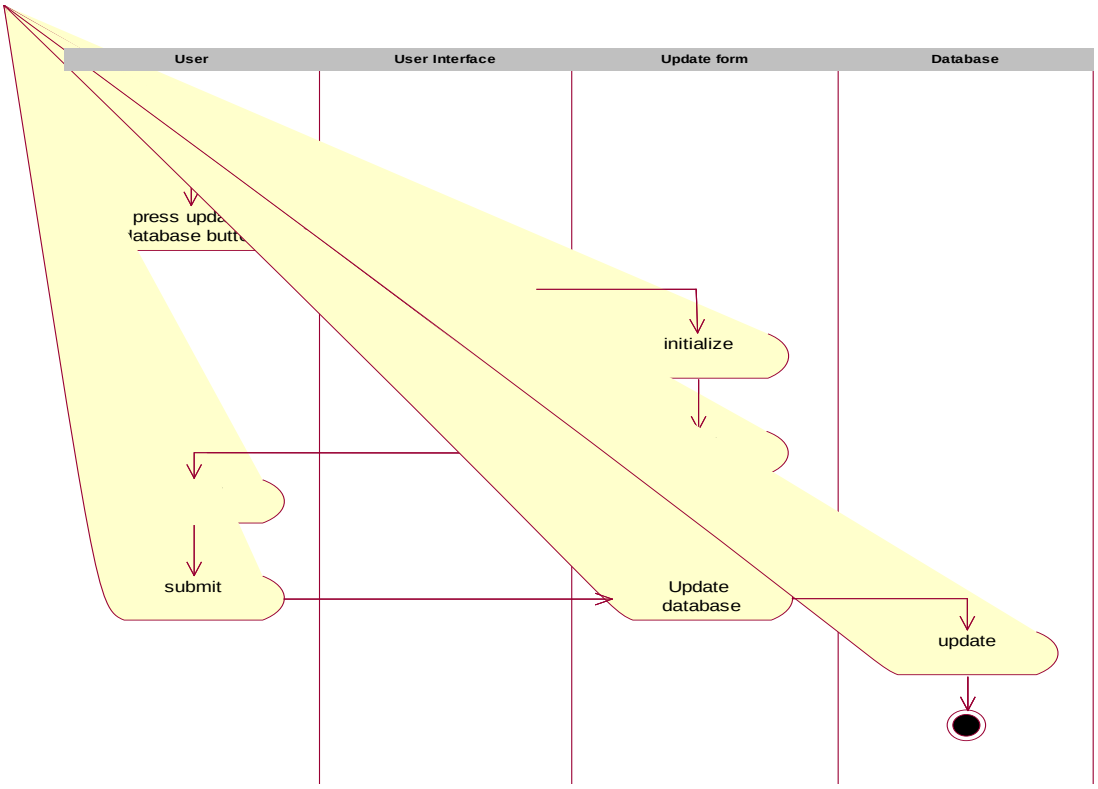
الشكل(٤-١٦) الأحداث لعملية المسح clear

المنارة للاستشارات
www.manaraa.com



الشكل(٤-١٧) الأحداث لعملية التزاوج الاختياري Custom mutation

www.manaraa.com



الشكل(٤-١٨) الأحداث لعملية تجديد قاعدة البيانات Update database

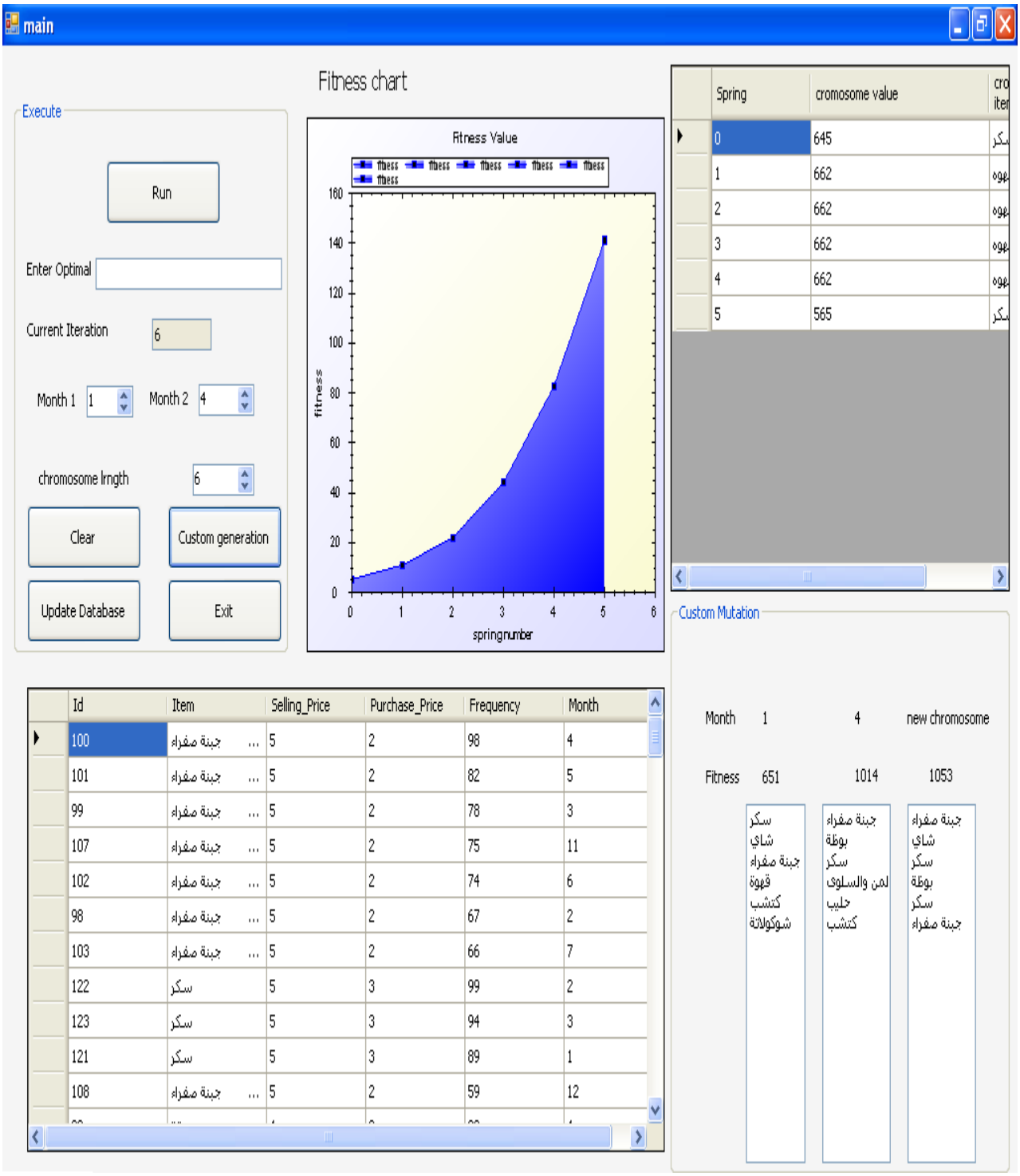
٤،٤ التطبيق Application

٤،٤،١ كيفية استخدام هذا البرنامج

يعتبر النظام من ناحية الاستخدام بسيط بحيث يسمح لأي شخص بإستخدامه بغض النظر عن معرفته بالبرمجة وخبرته بالحاسوب بما ان البيانات مرتبة في جداول والواجهة المستخدمة سهلة الإستخدام وبسيطة، الشكل(٤-١٩) يبين الواجهة الرئيسية للنظام.

المنارة للاستشارات
www.manaraa.com

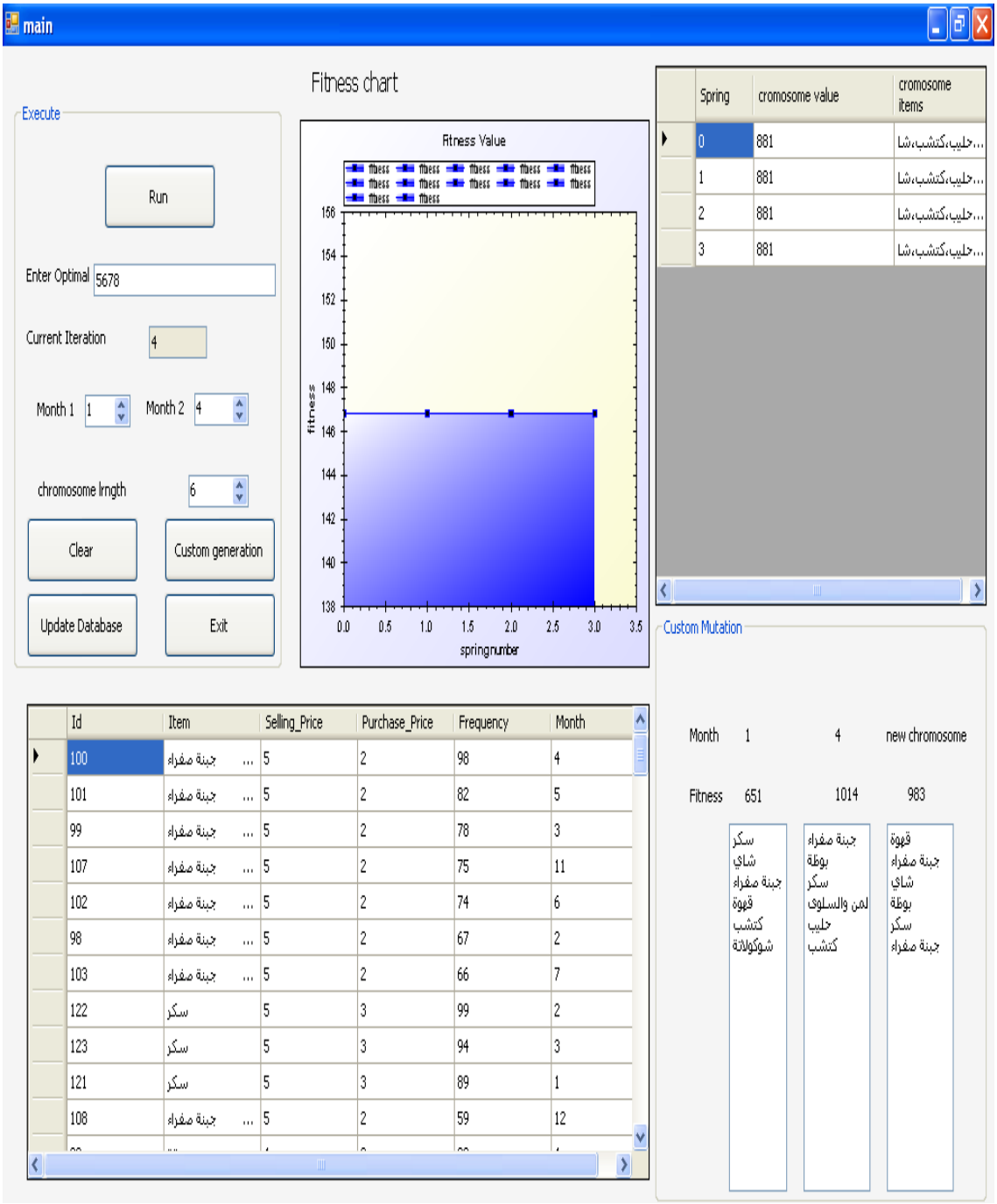
يبن الشكل (٤-١٩) الواجهة الرئيسية للنظام بعد تفعيل النظام.



الشكل(٤-١٩) الواجهة الرئيسية للنظام

المنارة للاستشارات
www.manaraa.com

عندما نحتاج ان نشغل البرنامج لنرى مستوى التحسن الذي طرأ على المجتمع البدائي نستطيع بسهولة ان نضغط على الزر RUN والشكل التالي يبين كيف أن البرنامج يطور مستوى الجودة للعناصر بما أن مقدار اللياقة Fitness يتزايد .



الشكل(٤-٢٠) ناتج عمليه التشغيل Run

المنارة للاستشارات
www.manaraa.com

أمـا زر CLEAR فيسـمح للمسـتخدم مـن مسـح كـل القـيم السـابقه والبـدء بالعمليـه من جديد.

أما Optimal فهو القيمه المراد الوصول اليها ان امكن.

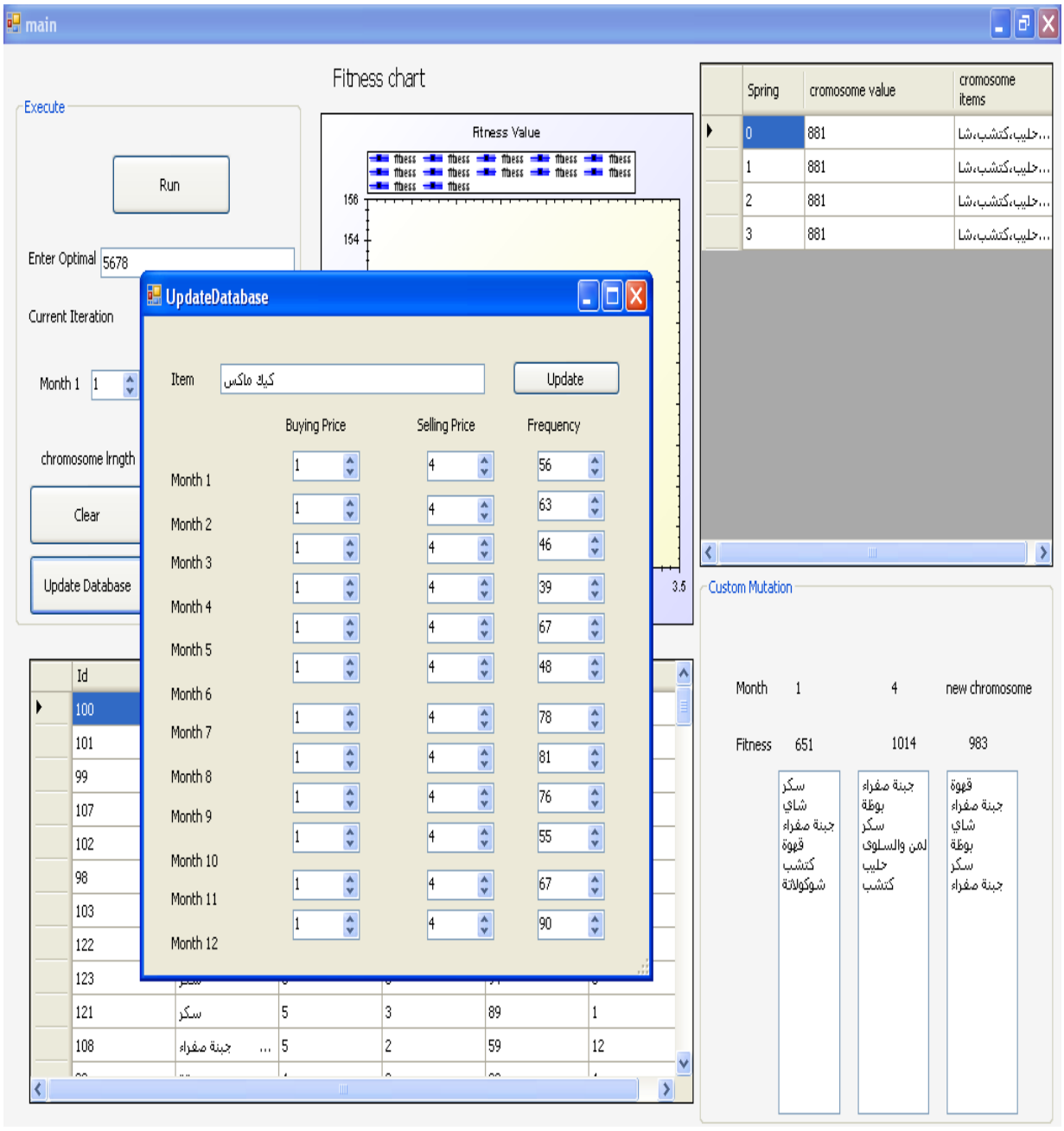
وبالنسـبة ل Current Iteration فهـذا الحقـل يظهـر عـدد الـدورات التـي قـام بهـا النظـام حتـى توقـف عـن المحاولـة للوصـول الـى الحـل المثـالي او الوصـول إلـى القيمـة المطلوبه، وليس لها أي اثر على تنفيذ البرنامج وهو يزيد كل مرة ينفذ بها البرنامج.

وأما Exit فتستخدم لإنهاء النظام.

ويمكننـا النظـام مـن القيـام بعمليـه تـزاوج اختياريـه بـين شـهرين مختلفـين لنـتمكن مـن روئيـه الكروموسـوم النـاتج عـن هـذا التـزاوج. وكمـا يسـمح للمسـتخدم مـن الـتحكم فـي طـول الكروموسـوم، وكمـا يسـمح النظـام بإضـافه سـلعه او تغييـر قيمهـا مـن خـلال شاشـه تجديد قاعدة البيانات.

المنارة للاستشارات
www.manaraa.com

يمثل الشكل (٤-٢١) ناتج عملية تجديد قاعدة البيانات update



الشكل (٤-٢١) ناتج عملية تجديد قاعدة البيانات update

المنارة للاستشارات
www.manaraa.com

الفصل الخامس

النتائج

المنارة للاستشارات
www.manaraa.com

٥.١ مقدمة:

نستعرض في هذا الفصل مزايا النظام المقترح ومدى اختلافه في التعامل مع الأنظمة الأخرى و ماهي الإضافات التي تمت في النظام المقترح،ثم نستعرض نتائج التنفيذ للنظام المقترح.

٥.٢ مزايا النظام المقترح:

يتميز النظام المقترح ببناء نظام جديد للتنقيب عن البيانات، وقد قمنا من خلال هذه الدراسة بتطوير عمليه تداخل بحيث تأخذ أفضل عناصر من كروموسومين، ويراعى فيها ان يتمكن المستخدم من التحكم في طول الكروموسوم وعملية التزاوج بين اشهر يحددها المستخدم،وايجاد عدد من الكروموسومات من خلال عملية التزاوج وظهور كروموسومات جديدة. ونقدم خوارزميه جينية مبنية لغايات تحسين اتخاذ القرارات.

وكما ان النظام المقترح اعتمد في بناء الكروموسومات على الاشهر خلال السنة بينما اتجهت الدراسات السابقة على الاعتماد على مبدا التزاوج بين عنصر وعنقود (Cluster) يمثل نوع من المواد التموينية.

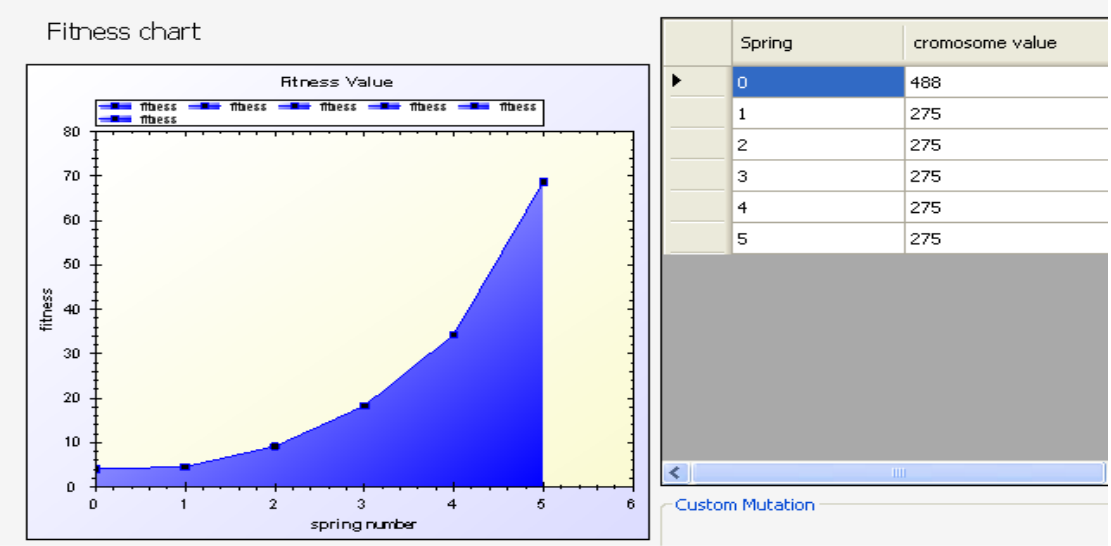
٥،٣ نتائج النظام

٥،٣،١ نتائج عملية تشغيل النظام

وفيما يلي بعض نتائج تشغيل النظام المقترح وتظهر في الشكل نتائج عملية التزاوج العشوائي وكيفية التحسن في مقدار اللياقة.

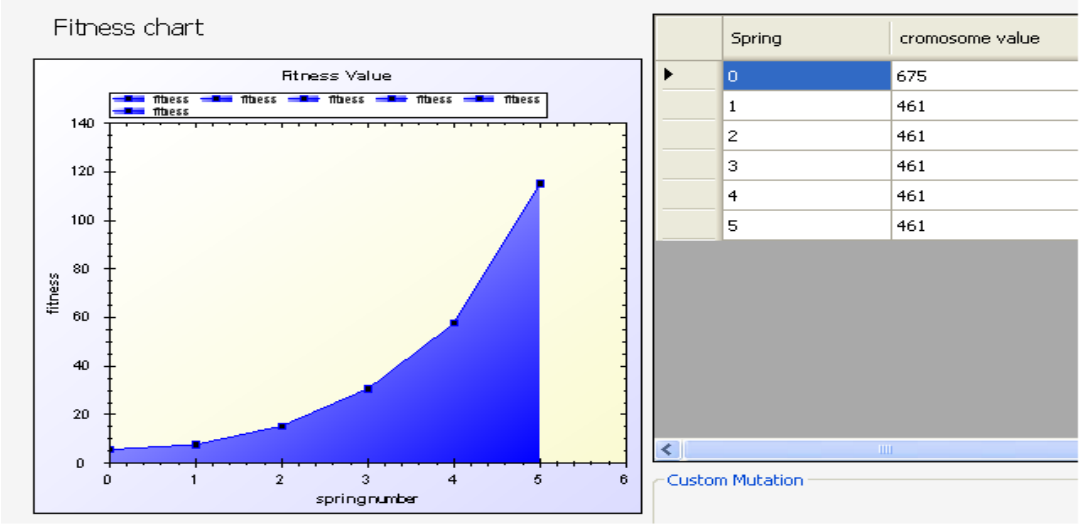
المنارة للاستشارات
www.manaraa.com

يبـين الشـكل (٥-١) إحـدى نتـائج تشـغيل النظـام حيـث يظهـر مـن خـلال الشـكل أن مقدار اللياقة يتزايد.



الشكل (٥-١) نتائج تشغيل النظام

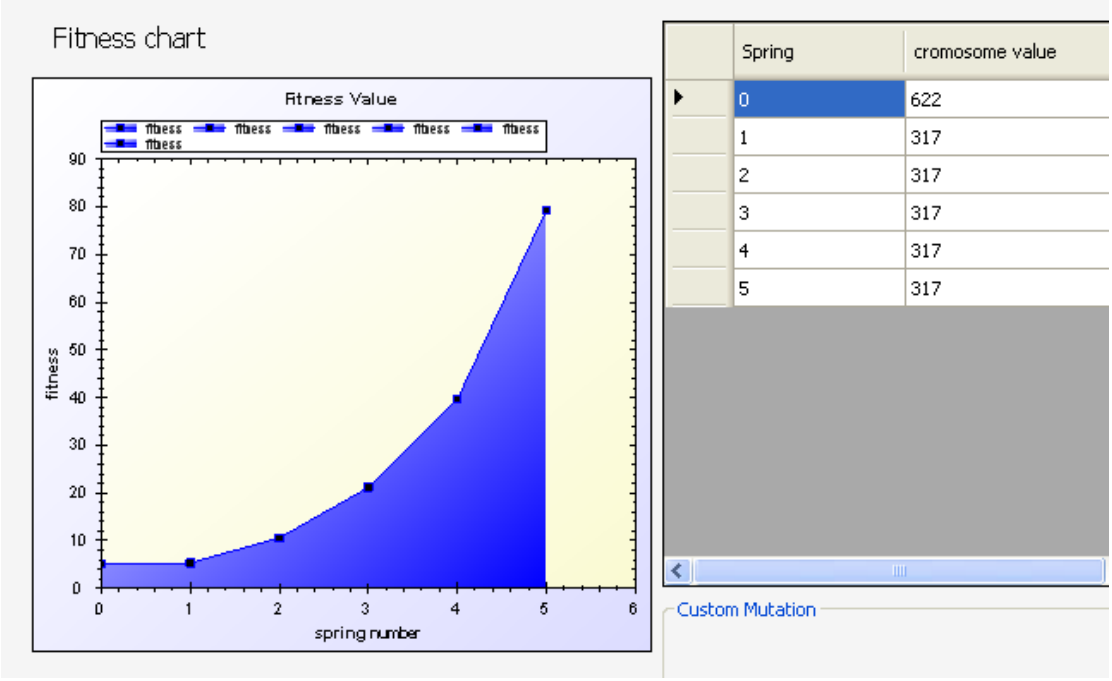
يبـين الشـكل (٥-٢) إحـدى نتـائج تشـغيل النظـام حيـث يظهـر مـن خـلال الشـكل أن مقدار اللياقة يتزايد.



الشكل(٥-٢) نتائج تشغيل النظام

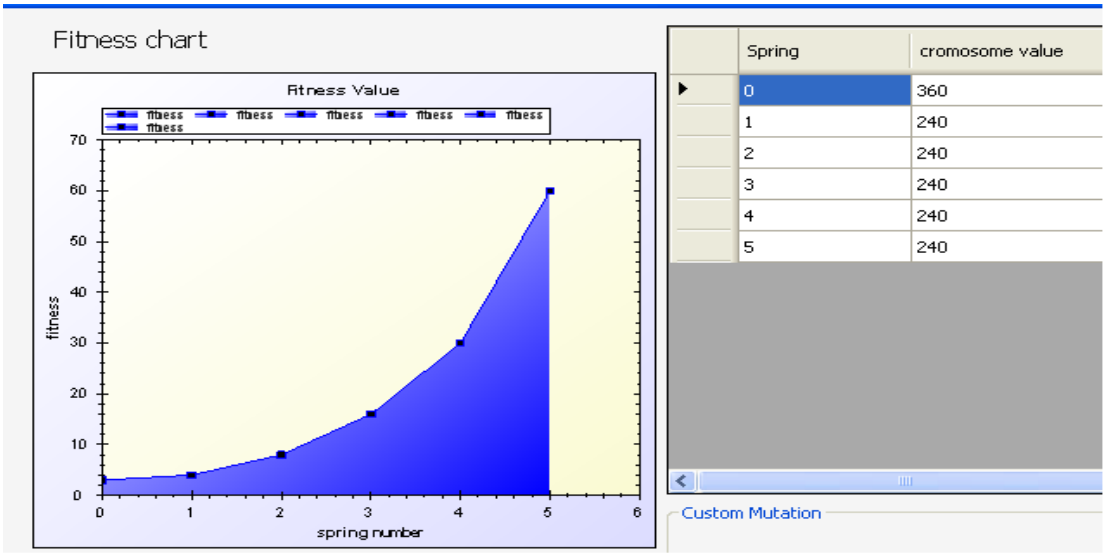
المنارة للاستشارات
www.manaraa.com

يبين الشكل (٥-٣) إحدى نتائج تشغيل النظام حيث يظهر من خلال الشكل أن مقدار اللياقة يتزايد.



الشكل(٥-٣) نتائج تشغيل النظام

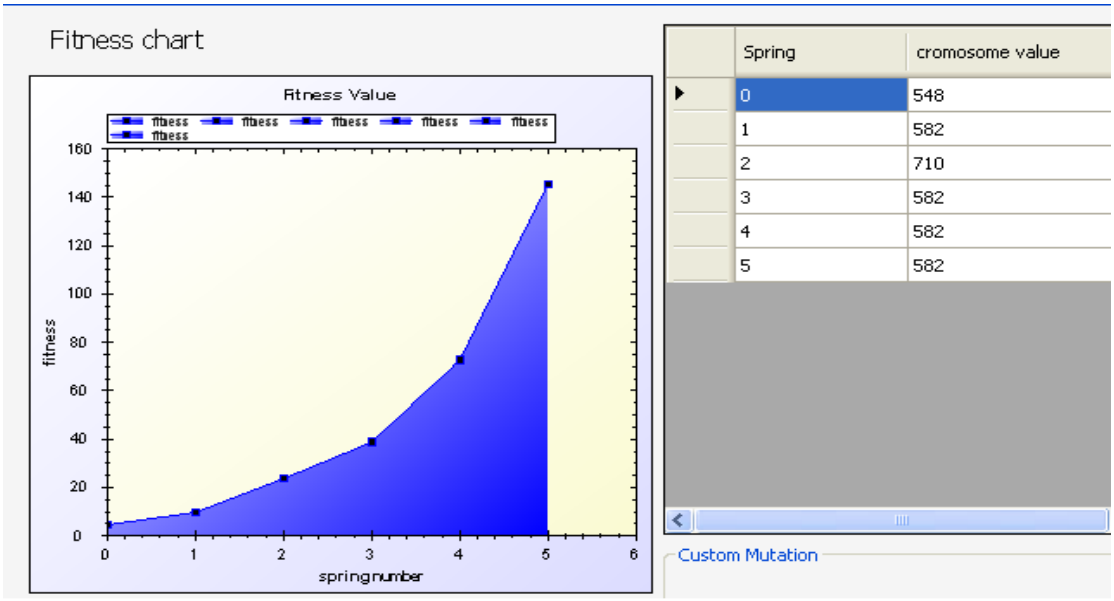
يبين الشكل (٥-٤) إحدى نتائج تشغيل النظام حيث يظهر من خلال الشكل أن مقدار اللياقة يتزايد.



الشكل(٥-٤) نتائج تشغيل النظام

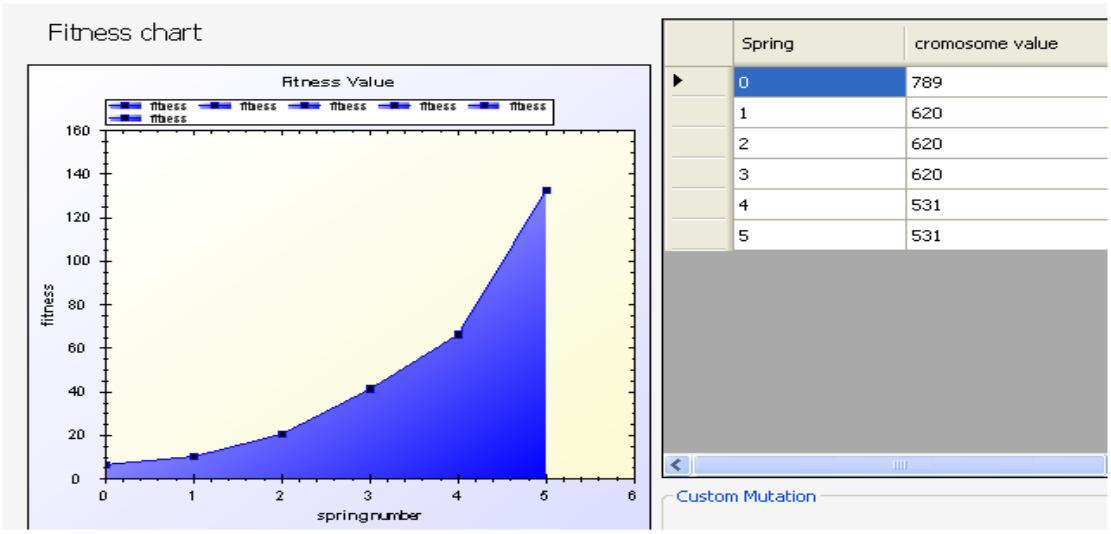
المنارة للاستشارات
www.manaraa.com

يبـين الشـكل (٥-٥) إحـدى نتـائج تشـغيل النظـام حيـث يظهـر مـن خـلال الشـكل أن مقدار اللياقة يتزايد.



الشكل(٥-٥) نتائج تشغيل النظام

يبـين الشـكل (٥-٦) إحـدى نتـائج تشـغيل النظـام حيـث يظهـر مـن خـلال الشـكل أن مقدار اللياقة يتزايد.



الشكل(٥-٦) نتائج تشغيل النظام

المنارة للاستشارات
www.manaraa.com

٥،٣،٢ نتائج عملية التزاوج الاختياري

يتم تمثيل الكروموسوم عن طريق مجموعه من أفضل العناصر (السلع) في شهر محدد بحيث يحتوي الكروموسوم على أفضل العناصر من ناحيه المردود وهو ممثل بالمعادله التاليه:

المردود = (سعر البيع – سعر التكلفة)*الكميه المباعة في ذلك الشهر

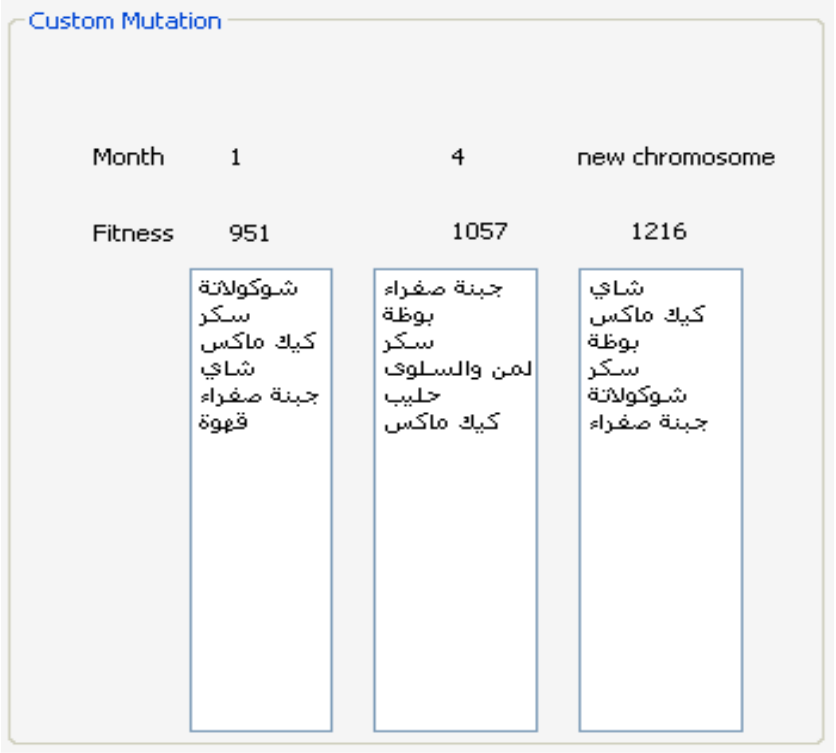
بحيث تكون العناصر مرتبه تنازليا من الشمال الى اليمين وبالنسبة لداله اللياقة (fitness) فتعبر عن مجموع المردودات من هذه العناصر، وفي المثال التالي توضيح لعمليه بناء الكروموسوم للشهر الأول

اسم السلعة	سعر الشراء	سعر البيع	الكميه المباعه	المردود
شوكولاتة	٢	٥	٩٠	٢٧٠
سكر	٣	٥	٨٩	١٧٨
كيك ماكس	١	٤	٥٦	١٦٨
شاي	١	٣	٦٥	١٣٠
جبنه صفراء	٢	٥	٣٥	١٠٥
قهوة	١	٢	١٠٠	١٠٠
مقدار اللياقة للشهر الأول				٩٥١

وفيما يلي بعض نتائج عملية التزاوج الاختياري لعدة أشهر.

يبين الشكل(٥-٧) عملية التزاوج بين الشهر الأول حيث يمتلك مقدار لياقة ٩٥١ والشهر الرابع حيث يمتلك مقدار لياقة ١٠٥٧،ومقدار اللياقة للكروموسوم الجديد ١٢١٦ وهي اعلي من مقدار اللياقة لكل من الأب وإلام.

المنارة للاستشارات
www.manaraa.com



الشكل(٥-٧)

نتائج عملية التزاوج الاختياري

يبين الشكل(٥-٨) عملية التزاوج بين الشهر الأول حيث يمتلك مقدار لياقة ٩٥١ والشهر السابع حيث يمتلك مقدار لياقة ٨١٥،ومقدار اللياقة للكروموسوم الجديد ١٩٨٤ وهي اعلي من مقدار اللياقة لكل من الأب وإلام.

الشكل(٥-٨)

نتائج عملية التزاوج الاختياري

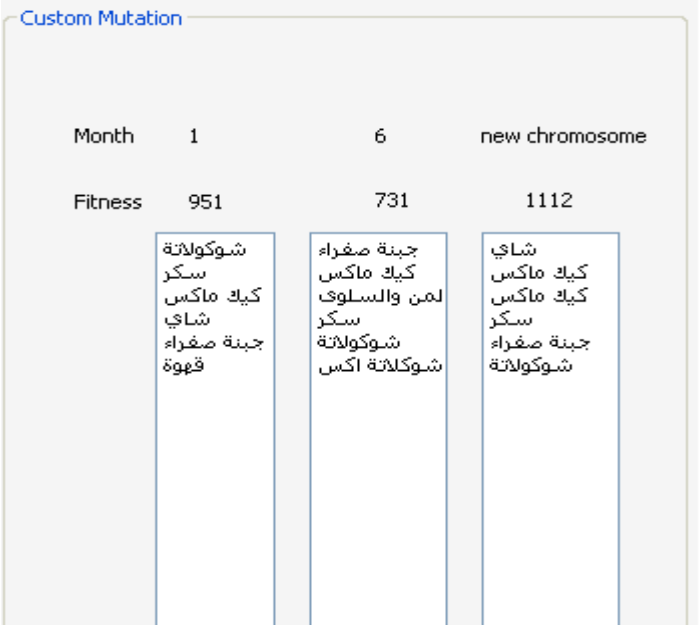
المنارة للاستشارات
www.manaraa.com

يبين الشكل(٥-٩) عملية التزاوج بين الشهر الأول حيث يمتلك مقدار لياقة ١١١٨ والشهر الخامس حيث يمتلك مقدار لياقة ١٠٥٣،ومقدار اللياقة للكروموسوم الجديد ١٣٦٠ وهي اعلي من مقدار اللياقة لكل من الأب وإلام.

الشكل(٥-٩) نتائج عملية التزاوج الاختياري

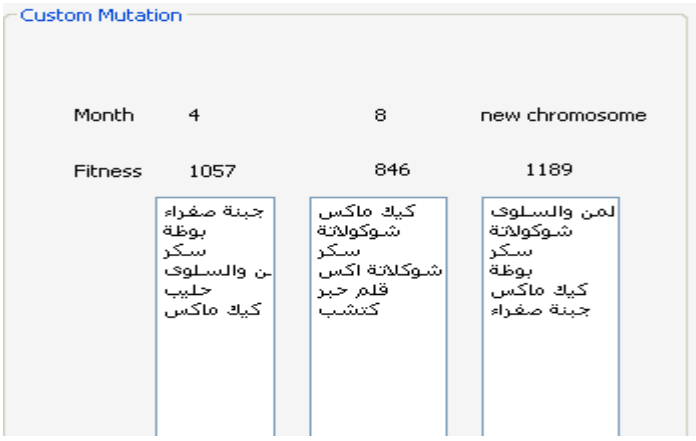
يبين الشكل(٥-١٠) عملية التزاوج بين الشهر الأول حيث يمتلك مقدار لياقة ٩٥١ والشهر السادس حيث يمتلك مقدار لياقة ٧٣١،ومقدار اللياقة للكروموسوم الجديد ١١١٢ وهي اعلي من مقدار اللياقة لكل من الأب وإلام.

المنارة للاستشارات
www.manaraa.com



الشكل(٥-١٠) نتائج عملية التزاوج الاختياري

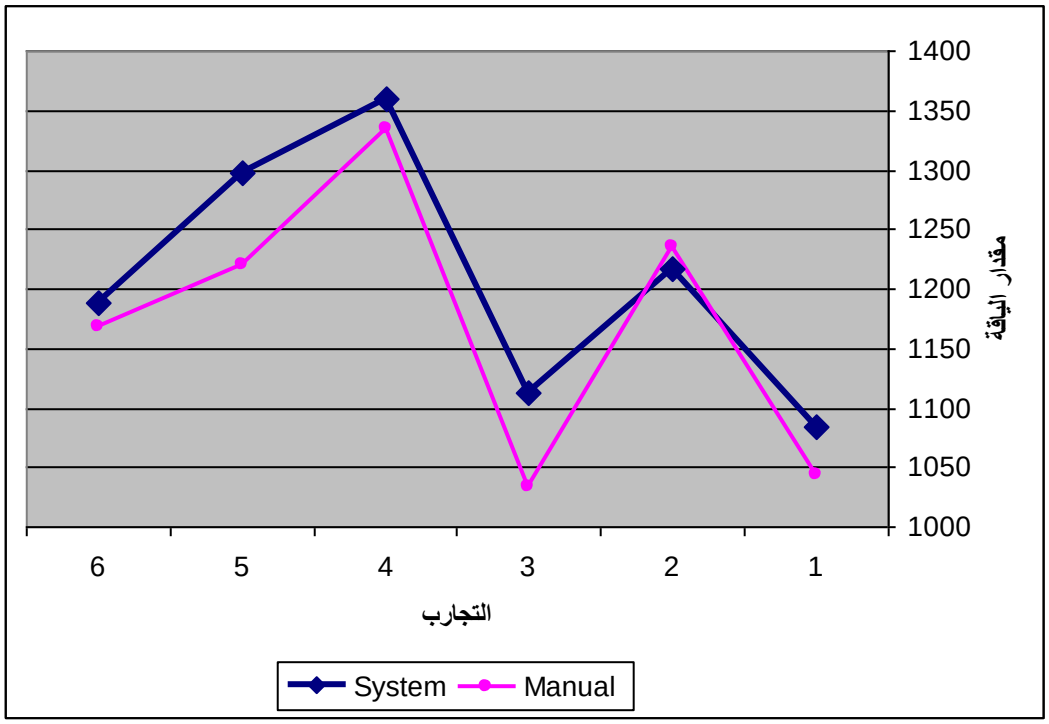
يبين الشكل(٥-١١) عملية التزاوج بين الشهر الرابع حيث يمتلك مقدار لياقة ١٠٥٧ والشهر الثامن حيث يمتلك مقدار لياقة ٨٤٦،ومقدار اللياقة للكروموسوم الجديد ١١٨٩ وهي اعلي من مقدار اللياقة لكل من الأب وإلام.



الشكل(٥-١١) نتائج عملية التزاوج الاختياري

المنارة للاستشارات
www.manaraa.com

يبن الشكل(٥-١٢) نتائج تنفيذ النظام المقترح ونتائج التنفيذ يدويا. حيث يبن الشكل أفضلية في الأداء لنتائج النظام المقترح على نتائج التنفيذ اليدوي.



الشكل(٥-١٢) مقارنة بين نتائج للنظام المقترح والتنفيذ اليدوي.

والجدول التالي يبين قيم الياقة الناتجة عن النظام المقترح والنظام اليدوي،ويتضح من خلال الجدول ان هناك تحسن في مقدار اللياقة لصالح النظام المقترح.حيث ان متوسط اللياقة الناتج للنظام المقترح هو١٢٠٩.٦٧ بينما متوسط اللياقة للنظام اليدوي هو١١٧٢.٥٠.

جدول(٥-١) يمثل نتائج التجارب للنظام اليدوي والنظام المقترح

Manual	System	No.exp
1044	1084	1
1236	1216	2
1033	1112	3
1334	1360	4
1220	1297	5
1168	1189	6

المنارة للاستشارات
www.manaraa.com

٥،٤ نتائج البحث والعمل المستقبلي

تم من خلال هذه الدراسة بناء نظام جديد للتنقيب عن البيانات، وقد قمنا من خلال هذه الدراسة بتطوير عمليه تداخل بحيث تأخذ أفضل عناصر من كروموسومين، ويراعى فيها ان يتمكن المستخدم من التحكم في طول الكروموسوم وعملية التزاوج بين أشهر يحددها المستخدم، وإيجاد عدد من الكروموسومات من خلال عملية التزاوج وظهور كروموسومات جديدة. ونقدم خوارزميه جينية مبنية لغايات تحسين اتخاذ القرارات. حيث ان النظام المقترح له أفضلية في الأداء مقارنة مع نتائج التنفيذ اليدوي.

من خلال هذا البحث تم التطرق لاستخدام الخوارزميات الجينية في عملية التنقيب عن البيانات كإحدى الخوارزميات التطويرية Evolutionary Algorithms ،ويمكن استخدام خوارزميات اخرى مثل خوارزمية النمل Ant Colony وهي خوارزمية تستخدم الاحتمالية لحل المسائل الحسابية التي يمكن ان تتحول لايجاد الطرق الجيدة من خلال الرسوم البيانية من خلال البحث عن الطريق المثالي من خلال الرسم البياني بطريقة مشابهه لسلوك النمل في ايجاد الطعام، او خوارزمية النحل Bee Colony وتعتمد هذه الخوارزمية على المسار الاولي المخزن مسبقا ومقارنة هذا المسار مع المسار المجاور وتحدد فيما اذا كان افضل او لا،وعندها يتم اعتماد المسار الأفضل، وتكرر هذه العملية الى ان يتم الحصول على افضل المسارات، والمقارنة مع الخوارزمية المستخدمة.

المنارة للاستشارات
www.manaraa.com

المصادر والمراجع

1. Belmont-Moreno, **The Role of Mutation and Population Size in Genetic Algorithms Applied to Physics Problems**, International Journal of Modern Physics C: Physics & Computers, Nov.2001, Vol. 12 Issue 9. PP1345-1356.
2. Binti Abdullah, Nik Nailah Liquiesre, Michel Cerri, Stefano, **A Ga' S Rule For Knowledge Discovery**, Applied Artificial Intelligence, May-Jul 2003, Vol 17, Issue 5. PP399-418.
3. Boutsinas, Basilis, **Accessing Data Mining Rules Through Expert Systems,** International Journal Of Information Technology & Decision Making, Dec.2002, Vol 1, Issue 4.PP657-673.
4. Coskun Samli, A.Pohlen , Terrance L.Bozovic,Nenad**, A Review Of Data Mining Techniques As They Apply To Marketing: Generating Strategic Information To Develop Market Segments**, Marketing Review, Winter 2002, Vol. 2 Issue 2. PP211-228.
5. Elian, Rich, **Artificial Intelligence**, second Edition, Person Education, 1991.
6. H.Chung, Michael Gray, Paul, **Special Section In Data Mining** , Journal Of Management Information Systems, Summer 99, Vol.16, Issue 1, PP11-17.
7. Jawei, han**, Data Mining concepts and Techniques**, forth edition, person Education, 2005.

المنارة للاستشارات

www.manaraa.com

8. Jeudy, Baptiste Boulicaut, Jean-François, **Optimization Of Association Rule Mining Queries**, Intelligent Data Analysis, 2002, Vol 6, Issue 4, PP341-358.

9. JR.Hruschka, R.Estevam, Hruschka, Eduardo, R.Ebecken, F.Nelson, A **Feature Selection Bayesian Approach For Extracting Classification Rules With A Clustering Genetic Algorithm**, Applied Artificial Intelligence, May-Jul2003, Vol 17, Issue 5, PP489-507.

10. Kumar, Vipin, **Introduction to Data Mining,** Pearson Addison, International Edition, 2006.

11. L. Chen, T. Sakaguchi And M. Frolick, **Data Mining Methods Applications And Tools**, Journal Of Information Systems Management, 2000, Vol 16, Issue 4, PP22-31.

12. Lorena, Luia Antonio Nogueira Furtado, J . Carlos, **Constructive Genetic Algorithm For Clustering Problem**, Evolutionary Computation, Fall 2001, Vol 9, Issue 3, PP309-328.

13. Lugar, George F, **Artificial Intelligence, Structures And Strategies For Complex Problem Solving**, Forth Edition, Pearson Education, 2002.

14. Mackinnon Murray, J.Glickned, **Data Mining And Knowledge Discovery In Database**, Australian And New Zealand Journal Of Statistics, Sep 2000, Vol 41, Issue 2, PP255-276.

15. Melab, Nordine**, Data Mining A Key Contribution To E-Business**, Information And Communication Technology Law, Oct2001, Vol 10, Issue 2, PP209-219.

المنارة للاستشارات

www.manaraa.com

16. Mustafa, Wael, **Optimization Of Production Systems Using Genetic Algorithms**, International Journal Of Computational Intelligence &Application, Sep 2003, Vol. 3 , Issue 3,PP233-249.

17. Negevitsky, Michael, **Artificial Intelligence, A Guide To Intelligent Systems, Second Edition**, Addison Wesley Education, 2005.

18. Ross, Peter, **What Are Genetic Algorithms Good** At, Journal On Computing; summer 97, Vol 9, Issue 3, PP260-263.

19. S.Poli, Riccardo Langdon, B.William,Schema **Theory For Genetic Programming With One-Point Crossover And Point Mutation**, Evolutionary Computation, Fall 98, Vol. 6, Issue 3, PP231-253.

20. Siddhartha Bhattacharyya, **Evolutionary Computation For Database Marketing**, Journal Of Database Marketing, 2002, Vol 10, Issue 4, PP242-252.

21. Yu-Ru Chen, I.Ming-Chuan Hung, I.Don-Lin Yang ,**Using Data Mining To Construct An Intelligent Web Search System,** International Journal Of Computer Processing Of Oriental Languages , Jun 2003,Vol 16 , Issue 2,PP143-171.

المنارة للاستشارات

www.manaraa.com

Program code:

```
Imports System.Data.SqlClient
Imports System.Data

Public Class Form1

    Private Sub Button3_Click(ByVal sender As
System.Object, ByVal e As System.EventArgs)
Handles Button3.Click
        Application.Exit()
    End Sub

    Private Sub Form1_Load(ByVal sender As
System.Object, ByVal e As System.EventArgs)
Handles MyBase.Load
        Dim i As Integer
        Dim j As Integer

        Randomize()
        Call Get_items()
        Current_Iteration = 0
        Cross_Over_Point = 3
        Mutation_Rate = 0.14
        optimal = 10000
        Population_Size = 12
        Curent_Population = 12
        For i = 0 To Population_Size - 1
            For j = 0 To 6
                week(i, j) = Int(Rnd() *
Items_Count) + 1
            Next
            Call Compute_Fitnes(i)
        Next
        Call print_pop()
    End Sub

    Dim optimal As Integer, Population_Size As
Integer
    Dim Curent_Population As Integer
    Dim week(0 To 12, 0 To 7) As Integer
    Dim Items(0 To 99) As Double
```

المنارة للاستشارات

www.manaraa.com

```
    Dim Items_Count As Integer
    Dim Cross_Over_Point As Integer
    Dim Mutation_Rate As Double
    Dim Current_Iteration As Integer

    Private Function Item_Profit(ByVal item As
Integer)
        Item_Profit = Items(item)
    End Function

    Private Sub Get_items()

        Dim con As SqlConnection = New
SqlConnection("Data
Source=.\SQLEXPRESS;AttachDbFilename='|DataDirect
ory|\dd.mdf';Integrated Security=True;User
Instance=True")
        Dim com As SqlCommand = New
SqlCommand("select * from Items", con)
        con.Open()
        Dim reader As SqlDataReader
        reader = com.ExecuteReader()



        Dim i As Integer = 0
        While reader.Read()
            Items(i) = (reader.GetDouble(2) -
reader.GetDouble(3))
            i = i + 1

        End While

        Items_Count = i

    End Sub

    Private Sub Compute_Fitnes(ByVal n As
Integer)
        week(n, 7) = 0
        Dim i As Integer

        For i = 0 To 6
```

المنارة للاستشارات
www.manaraa.com

```
            week(n, 7) = week(n, 7) +
Item_Profit(week(n, i)) ^ 2 + 10
        Next
    End Sub

    Private Sub print_pop()
        Dim i As Integer
        Dim j As Integer

        For i = 0 To Curent_Population
            For j = 0 To 7
                Pic.Text += week(i, j).ToString()

            Next
            Pic.Text += " "

        Next
        Pic.Text += Curent_Population.ToString()


    End Sub

    Private Sub kill_none_optimal()
        Dim avg As Double
        Dim i As Integer

        For i = 0 To Curent_Population - 1
            avg = avg + (week(i, 7) /
Curent_Population)
        Next
        i = 0
        Dim k As Integer
        Dim j As Integer

        While i <= Curent_Population - 1
            If week(i, 7) < avg Then 'kill
                For j = i To Population_Size - 1
                    For k = 0 To 7
                        week(j, k) = week(j + 1,
k)
                    Next k
                Next j
```

www.manaraa.com

```
                Curent_Population =
Curent_Population - 1
            Else
                i = i + 1
            End If
        End While
    End Sub

    Private Sub Make_Child()
        Dim ch1 As Integer, ch2 As Integer,
new_ch As Integer
        Dim i As Integer

        ch1 = Int(Rnd() * Curent_Population) + 1

        ch2 = Int(Rnd() * Curent_Population) + 1

        new_ch = Curent_Population + 1
        For i = 0 To 6
            If Rnd() > Mutation_Rate Then
                If i < Cross_Over_Point Then
                    week(new_ch, i) = week(ch1,
i)
                Else
                    week(new_ch, i) = week(ch2,
i)
                End If
            Else
                week(new_ch, i) = Int(Rnd() *
Items_Count) + 1
            End If

        Next

        Call Compute_Fitnes(new_ch)
        Curent_Population = Curent_Population + 1
    End Sub



    Private Sub Command1_Click()
        Pic.Clear()

    End Sub
```

المنارة للاستشارات
www.manaraa.com

```
    Private Sub Button1_Click(ByVal sender As
System.Object, ByVal e As System.EventArgs)
Handles Button1.Click
        Call kill_none_optimal()
        If Curent_Population <> Population_Size
Then

            While Curent_Population <
Population_Size
                Call Make_Child()
            End While

            Call print_pop()
            Current_Iteration = Current_Iteration
+ 1
            TextBox2.Text = Current_Iteration
        Else
            TextBox1.Text = ""
            Dim j As Integer

            For j = 0 To 6
                TextBox1.Text = TextBox1.Text +
Str(week(0, j))
            Next
            TextBox1.Text = TextBox1.Text + "
Fitness " + Str(week(0, 7))
        End If
    End Sub
End Class
```

www.manaraa.com

ملحق بتعليمات النظام

Program code:

```
Imports System.Data.SqlClient
Imports System.Data

Public Class Form1

    Private Sub Button3_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button3.Click
        Application.Exit()
    End Sub

    Private Sub Form1_Load(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles MyBase.Load
        Dim i As Integer
        Dim j As Integer

        Randomize()
        Call Get_items()
        Current_Iteration = 0
        Cross_Over_Point = 3
        Mutation_Rate = 0.14
        optimal = 10000
        Population_Size = 12
        Curent_Population = 12
        For i = 0 To Population_Size - 1
            For j = 0 To 6
                week(i, j) = Int(Rnd() * Items_Count) + 1
            Next
            Call Compute_Fitnes(i(
        Next
        Call print_pop()
    End Sub

    Dim optimal As Integer, Population_Size As Integer
    Dim Curent_Population As Integer
    Dim week(0 To 12, 0 To 7) As Integer
    Dim Items(0 To 99) As Double

    Dim Items_Count As Integer
    Dim Cross_Over_Point As Integer
    Dim Mutation_Rate As Double
    Dim Current_Iteration As Integer

    Private Function Item_Profit(ByVal item As Integer(
        Item_Profit = Items(item(
    End Function

    Private Sub Get_items()
```

المنارة للاستشارات

www.manaraa.com

```
        Dim con As SqlConnection = New SqlConnection("Data
Source=.\SQLEXPRESS;AttachDbFilename='|DataDirectory|\dd.mdf';Integrated
Security=True;User Instance=True("
        Dim com As SqlCommand = New SqlCommand("select * from Items", con(
        con.Open()
        Dim reader As SqlDataReader
        reader = com.ExecuteReader()



        Dim i As Integer = 0
        While reader.Read()
            Items(i) = (reader.GetDouble(2) - reader.GetDouble(3((
            i = i + 1

        End While

        Items_Count = i

    End Sub

    Private Sub Compute_Fitnes(ByVal n As Integer(
        week(n, 7) = 0
        Dim i As Integer

        For i = 0 To 6
            week(n, 7) = week(n, 7) + Item_Profit(week(n, i)) ^ 2 + 10
        Next
    End Sub

    Private Sub print_pop()
        Dim i As Integer
        Dim j As Integer

        For i = 0 To Curent_Population
            For j = 0 To 7
                Pic.Text += week(i, j).ToString()

            Next
            Pic.Text" " =+

        Next
        Pic.Text += Curent_Population.ToString()


    End Sub

    Private Sub kill_none_optimal()
        Dim avg As Double
        Dim i As Integer
```

www.manaraa.com

```
    For i = 0 To Curent_Population - 1
       avg = avg + (week(i, 7) / Curent_Population(
    Next
    i = 0
    Dim k As Integer
    Dim j As Integer

    While i <= Curent_Population - 1
       If week(i, 7) < avg Then 'kill
          For j = i To Population_Size - 1
             For k = 0 To 7
                week(j, k) = week(j + 1, k(
             Next k
          Next j

          Curent_Population = Curent_Population - 1
       Else
          i = i + 1
       End If
    End While
End Sub

Private Sub Make_Child()
    Dim ch1 As Integer, ch2 As Integer, new_ch As Integer
    Dim i As Integer

    ch1 = Int(Rnd() * Curent_Population) + 1

    ch2 = Int(Rnd() * Curent_Population) + 1

    new_ch = Curent_Population + 1
    For i = 0 To 6
       If Rnd() > Mutation_Rate Then
          If i < Cross_Over_Point Then
             week(new_ch, i) = week(ch1, i(
          Else
             week(new_ch, i) = week(ch2, i(
          End If
       Else
          week(new_ch, i) = Int(Rnd() * Items_Count) + 1
       End If

    Next

    Call Compute_Fitnes(new_ch(
    Curent_Population = Curent_Population + 1
End Sub
```

www.manaraa.com

```
    Private Sub Command1_Click()
        Pic.Clear()

    End Sub


    Private Sub Button1_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles Button1.Click
        Call kill_none_optimal()
        If Curent_Population <> Population_Size Then

            While Curent_Population < Population_Size
                Call Make_Child()
            End While

            Call print_pop()
            Current_Iteration = Current_Iteration + 1
            TextBox2.Text = Current_Iteration
        Else
            TextBox1.Text"" =
            Dim j As Integer

            For j = 0 To 6
                TextBox1.Text = TextBox1.Text + Str(week(0, j((
            Next
            TextBox1.Text = TextBox1.Text + " Fitness " + Str(week(0, 7((
        End If
    End Sub
End Class
```

www.manaraa.com

Abstract

Data mining is considered as one of the most important fields in information technology, as information enormously increases, and for the difficulty of digging for important data of this quantity. Data mining became increasingly used in many fields, i.e., credit cards, banks, marketing, industry, telephone companies, and many others. In general, mining is divided into two methods: using statistics and second most important method used in artificial intelligence is machine learning (Siddhartha, 2002).

This study will use genetic algorithms in the process of data extraction and mining from a data base for a supply center. There are some characteristics that make genetic algorithms suitable for difficult problems, as they are designed to provide the best performance with unwritten variables with the data containing errors, in addition to its great ability on searching, in a great group of examples limited with specific rules, for the best solution (Mustafa, 2003; Binti Abdullah, et al).

The proposed system is distinguished with a new data mining structure. This study developed the crossover, where the best elements are taken from two chromosomes, taking into consideration that the user has control in the length of the chromosome, and the custom mutation between months controlled by the user, and finding the number of chromosomes from mating process and the emergence of new chromosomes. The study provides a new genetic algorithm structured to improve decision making.

www.manaraa.com