بسم الله الرحمن الرحيم

جامعة السودان للعلوم و التكنولوجيا

كلية علوم الحاسوب وتقانة المعلومات

قسم هندسة البرمجيات

نظام خبير هجيني لتشخيص اورام الثدي

مشروع مقدم كأحد متطلبات الحصول على بكالريوس الشرف في هندسة البرمجيات

اعداد الطالبات :

سماح اسحاق ادم

هالة يوسف محمد

توقيع المشرف

د/ هويدا علي عبدالقادر

اغسطس 2014

بسم الله الرحمن الرحيم

جامعة السودان للعلوم و التكنولوجيا

كلية علوم الحاسوب وتقانة المعلومات

قسم هندسة البرمجيات

نظام خبير هجيني لتشخيص اورام الثدي

مشروع مقدم كأحد متطلبات الحصول على بكالريوس الشرف في هندسة البرمجيات

اعداد الطالبات :

سماح اسحاق ادم

هالة يوسف محمد

توقيع المشرف

د/ هويدا علي عبدالقادر

.....

التاريخ : / /

الآيــة

قال تعالى :

(وقل ربي زدني علما)

صدق الله العظيم

سورة طه الآية (114)

الَاهـــــــداء

الى من كللهم الله بالهيبة و الوقار

الى من نحمل أسمائهم بكل افتخار

"أبـــــــــائنا الاعـــــــزاء "

الى من كانت دعواتهم اسرار نجاحنا

الى من كان حنانهن بلسم جراحنا

الى من كان كفاحهن سر حياتنا

"امهـــــــاتنا الحبيبـــــــات "

الى من تذوقنا معهم أجمل اللحظات

الى من نحبهم حبا لو مر على ارض قاحلة لتفجرت منها ينابيع المحبة متدفقات

"الاخـــــــــوة و الاخـــــــــــوات "

الى من سرنا سويا و نحن نشق طريق النجاح و الابداع

"الزمـــــــــــــلاء و الزمـــــــــــــيلات "

الى الشموع التي تحترق لتضئ للاخرين

"الاســـــــــــــاتذة الاجـــــــــــــلاء "

الى من كانوا يضيئون لنا الطريق و يساندوننا في السراء و الضراء

"الاصـــــــــدقاء الوفـــــــــــياء "

الى من اجلنا ذكره ليكون مسك الختام

مشرفتنا الدكتورة / هويدا علي عبدالقادر

لك نهدي هذا البحث المتواضع راجين من المولى عز و جل ان يجد القبول و النجاح.

الشكـــر و العـــرفان

أولا و أخيرا الشكر لله وحده الذي وفقنا و أعاننا على انجاز هذا البحث فله الحمد و الشكر و الثناء .

ثم نتقدم بالشكر و العرفان لأمهاتنا و أبائنا و الى زوجي العزيز / خليل جبارة الله بارك الله فيهم و عافاهم فهم من اسدوا الينا جميل المعروف .بفضلهم وصلنا الى اتمام هذه الرسالة جزاهم الله خير الجزاء و اعاننا على برهم و الاحسان اليهم .

كما نتقدم بالشكر و العرفان لاعضاء هيئة التدريس و كل من ساهم معنا في هذا البحث .

المستخلص

مع التقدم التكنلوجي الذي يشهده عصر العولمة بات على المجتمع أن يواجه و يواكب التطور التكنلوجي بفهم و عناية و متابعة , و بما ان الصحة هي العنصر المهم و الاول للبشرية ويؤثر عليها في مختلف اتجاهاتها لذلك لابد من تقنية تشخيص طبية لمرض اورام الثدي الذي يصيب و يوثر على شريحة حيوية و كبيرة من المجتمع .

تكمن مشكلة البحث في عدم توفر الخبرات البشرية بصورة دائمة , كما يمكن حدوث أخطاء تشخيصية يغفل عنها و تزداد احتمالية الخطأ كلما قلت خبرة الطبيب.

لقد قام الباحثون بدراسة الوضع الراهن وتم اقتراح حل وهو تصميم نظام خبير هجيني باستخدام خوازمية تنقيب البيانات و تقنية التظم الخبيرة باستخدام برنامج Exsys Corvid و Clementine لاستفادة من مخرجات التنقيب في البيانات و التي تمثل قوانين ارتباط و علاقات تساعد في التشخيص , لتصبح قاعدة معرفة للنظام الخبير .

من اهم النتائج التي توصل اليها البحث مساعدة الاطباء غير ذوي الخبرة الكافية في تحديد نوع الورم من خلال المعطيات و القراءات التي تختبر بواسكة الكشف عن أورام الثدي .

يوصي الباحثون بتطبيق النظام داخل مراكز الكشف عن اورام الثدي و المستشفيات .

Abstarct

With technological advance and development of the area of globalization on the community to face and keep pace with evolving technological understanding and care and follow-up , and including that health is an important factor and the first of humaneness and efforts them in different tumors , which infects and affects the slice vitality and great of society .

Enable the research problem in the lack of human expertise are always , as can the occurrence of diagnostic errors overlooks and increases the endurance of error ,the less experienced doctor .

The researchers have studies the current situation and propose a solution has been modeled on the design of hybrid expert system technique is Exsys Corvid which provides a convenient development environment for building expert systems .

Of the most important finding of the research to help determine the type of tumor through the data and readings by testing device to detect breast tumors .

Researchers recommend the application of the system within the centers for detecting breast tumors , and hospitals.

فهرس المحتويات

الرقم	الموضوع	رقم الصفحة
1	الآية	ا
2	الاهداء	ب
3	الشكـــر و العرفان	ج
4	المستخلص	د
5	Abstract	ه
6	فهرس المحتويات	و
7	فهرس الاشكال	ط
8	فهرس الجداول	ي
الباب الأول		
1.1	تمهيد	**1**
2.1	مقدمة البحث	**1**
3.1	مشكلة البحث	**2**
4.1	الحل المقترح	**2**
5.1	اهداف البحث	**2**
1.5.1	اهداف عامة	**2**
2.5.1	اهداف خاصة	**2**
6.1	حدود البحث	**2**
7.1	أهمية البحث	**3**
8.1	المساهمات المتوقعة من البحث	**3**
الباب الثاني		

1.2	التنقيب عن البيانات	**4**
1.1.2	تعريف التنقيب عن البيانات	**4**
2.1.2	اسس التنقيب عن البيانات	**5**
3.1.2	محددات تنقيب البيانات	**5**
4.1.2	مجال تنقيب البيانات	**5**
5.1.2	التقنيات الاكثر استخداما في التنقيب عن اليبانات	**6**
6.1.2	أهداف تقيب البيانات	**6**
7.1.2	معمارية تنقيب البيانات	**7**
8.1.2	تنقيب البيانات و مناجم البيانات	**7**
2.2	النظم الذكية	**8**
1.2.2	تعريف المعرفة	**8**
2.2.2	تعريف الخبير بصفة عامة	**8**
3.2.2	تفكير الخبراء	**8**
4.2.2	القواعد كاسلوب لتمثيل المعرفة	**8**
5.2.2	اللاعبون الرئسيون في فريق تطوير الخبرة	**9**
1.5.2.2	أعضاء فريق تطوير النظام الخبير	**9**
2.5.2.2	أجزاء نظام الخبرة	**9**
6.2.2	الخواص الاساسية لنظام الخبرة	**9**
1.6.2.2	احتمالية خطأ النظم الخبيرة	**10**
7.2.2	أساليب استدلال التسلسل للامام و التسلسل للخلف	**10**
1.7.2.2	التسلسل للامام	**10**

2.7.2.2	التسلسل للخلف	**10**
3.2	النظم الهجين	**11**
1.3.2	تعريف النظم الهجين	**11**
2.3.2	الحوسبة اللينة	**11**
3.3.2	الفرق بين الحوسبة اللينة و الذكاء الاصطناعي	**11**
4.3.2	مكونات النظام الخبير الهجيني	**11**
5.3.2	نظم الخبرة العصبية	**12**
6.3.2	التفكير التقريبي	**12**
4.2	الدراسات السابقة	**13**
1.4.2	مرشد طبي خبير متعدد الوسائط	**13**
2.4.2	استخدام تقنية التنقيب في البيانات لمرضى القلب	**13**
الباب الثالث		
1.3	المقدمة	**14**
2.3	الادوات المستخدمة	**14**
1.2.3	Microsoft excel	**14**
2.2.3	Clementine	**14**
3.2.3	ExsysCorvid	**15**
4.2.3	HTML	**15**
3.3	التحليل	**16**
1.3.3	استخدام الكلمنتاين	**16**
1.1.3.3	خوارزمية C4.5	**16**
1.1.1.3.3	مقدمة	**16**

2.1.1.3.3	See5/C5.0	**17**
2.3.3	بناء النظم الخبيرة	**20**
الباب الرابع		
1.4	مقدمة	**33**
2.4	النتائج	**33**
3.4	التوصيات	**33**
4.4	الخاتمة	**34**
5.4	الملاحق	**35**
6.4	المراجع	**39**

فهرس الاشكال

رقم الشكل	اسم الشكل	رقم الصفحة
الباب الثاني		
1.2	المعرفة المتنوعة في تنقيب البيانات	**6**
الباب الثالث		
1.3	واجهة برنامج كلمنتاين	**17**
2.3	القوانين الناتجة من خوازمية C5	**18**
3.3	تحليل القواعد	**19**
4.3	interface لنظام خبير	**20**
5.3	مقدمة ExsysCorvid	**21**
6.3	شاشة الصلاحية فيExsysCorvid	**21**

7.3	ضبط ExsysCorvid	**22**
8.3	ضبط TomCat	**22**
9.3	الشاشة الرئيسية في ExsysCorvid	**23**
10.3	إنشاء المتغيرات	**24**
11.3	إنشاء الاوامر المنطقية للمتغيرات	**24**
12.3	إنشاء الاوامر المنطقية للقواعد	**25**
13.3	إنشاء اوامر الترتيب	**25**
14.3	ادخال اسم المريض	**26**
15.3	ادخال عمر المريض	**26**
16.3	تحديد جنس المريض	**27**
17.3	ادخال قراءة 10	**27**
18.3	ادخال قراءة AREA	**28**
19.3	ادخال قراءة DA	**28**
20.3	ادخال قراءة A/DA	**29**
21.3	ادخال قراءة DR	**29**
22.3	ادخال قراءة HFS	**30**
23.3	ادخال قراءة P	**30**
24.3	ادخال قراءة MAXIP	**31**
25.3	ادخال قراءة PA500	**31**
26.3	مثال لتنفيذ البرنامج	**32**

فهرس الجداول

رقم الجدول	اسم الجدول	رقم الصفحة
الباب الرابع		
1.5	Data Set	36

البـــــاب الأول

المقدمة

1.1 تمهيد:

سرطان الثدي (Breast Cancer) بالنسبة للإناث هو أهم وأخطر السرطانات من حيث المكان والوفيات وكذلك معدل الإصابة، حيث يعتبر أكثر السرطانات شيوعا في العالموتشير الدراسات إلى أن امرأة واحدة من بين كل ثماني سيدات معرضة للإصابة ب سرطان الثدي في فترة ما من حياتها.

وتصاب أكثر من مليون امرأة سنوياًحول العالم بمرض سرطان الثدي. وتزيل الجراحة الورم بينما تخضع المريضة بعد ذلك للعلاج بالأشعة لقتل الخلايا الكامنة في الثدي .
سئكما أن سرطان الثدي يمثل أفضل نموذج لورم سرطاني تتغير فيه طبيعة العلاج ونتائجه طبقا لمرحلة الاكتشاف (مبكراً كان أو متأخراً) كما أن توعية عامة الناس لها تأثير ايجابي على السيطرة على المرض, والشفاء منه وكونه لا توجد وسيلة لتفادي الإصابة بهذا المرض، فإن تركيز الأطباء في تخفيض الوفيات منه هو التشخيص المبكر .

والتشخيص المبكر يعني أن السرطان محصور في منطقة ضيقة ولم ينتشر إلى أعضاء حيوية أخرى، ويمكن استئصاله بجراحة ومواصلة العلاجين الكيميائي والإشعاعي والخضوع لمراقبة دورية روتينية.

2.1 مقدمة البحث :

يعتمد الطبيب في تشخيصه لسرطان الثدي على الفحص
الفيزيائي(السريري)والتصويرالإشعاعي (الماموغرام) بحيث يراقب أي تغيرات في شكل الثدي أو وجود كتل تحت الجلد أو وجود إفرازات حليبية مائلة للصفرة، ويتم تأكيد الإصابة غالباً بسحب عينةنسيجية من الثدي ودراستها مخبرياًجاءت فكرة الاستفادة من التقنيات الحديثة في البيئة الطبية و التشخيصية لتساعد الأطباء في تشخيص المرض واكتشاف نوعه .

في حالات الكشف المبكر قد يكون الورم الذي يتم اكتشافه من أنواع الأورام الخبيثة في أقل من 20%من الحالات.

مع التطور السريع في حقول الطب وعلوم الحاسب، أصبح ما يسمى بالطبيب الإلكتروني واقعاً ملموساً يمكن الوثوق به وعاملاً مساعداً في الكشف عن الكثير من الأمراض وتشخيصها بشكل مبكر مما يساعد على الشفاء المبكر أو حتى تجنب الكثير من هذه الأمراض. ويتنامى دور هذه النظم مع التقدم الكبير في طرق التشخيص الطبي التي تنتج كميات هائلة من المعطيات التي تحتاج لمن يفسرها ويستخلص منها معلومات تشخيص الأمراض بشكل سريع وفعّال.

نظام تشخيص أو تحديد الأمراض بمساعدة الحاسب Computer-AidedDetection/Diagnosis – CAD، هو عبارة عن تقنية تساعد أطباء الأشعة في تفسير الصور الطبية بشكل دقيق وتحديد الأعراض المحتملة وتجنب التفسير الخاطئ أو إهمال بعض

الآفات المرضية نتيجة التفسير الشخصي. وتجدر الإشارة بأن نظم التشخيص تقدم حلاً أو رأياً إضافيًا ولا يمكن أن يحل بديلاً عن طبيب الأشعة، الذي له الرأي الفصل في تحديد المرض' تستخدم نظم التشخيص بمساعدة الحاسب في الكشف المبكر عن الكثير من الأمراض مثل سرطان الثدي، حيث يخفض استخدام نظم التشخيص حالات الوفيات الناتجة عن سرطان الثدي, كما تستخدم في الكشف عن العديد من أمراض القلب والأوعية الدموية، وفي الكشف عن الأورام التي قد تصيب مختلف أعضاء الجسم البشري، وغيرها الكثير من الأمراض.

3.1 مشكلة البحث :

- عدم توفر الخبرات البشرية بصورة دائمة.
- الخطأ الطبي في التشخيص هو نتيجة غير مرغوب فيها كان من الممكن تقليلها.
- قلة المعلومات الطبية لدى الطبيب المعالج.
- عدم توفر الإمكانيات التشخيصية أو لخطأ بالتحاليل أو الإجراءات الطبية.
- عدم كفاءة الأنظمة والإجراءات المستخدمة.

4.1 الحل المقترح للمشكلة :

يقدم البحث المقترح تصميماً جديداً مبتكراًلدمجخواص نظامين ذكيين لإخراج قوانين لبناء نظم ذكية هجينة لتشخيص أورام الثدي .

5.1 أهداف البحث :

1.5.1 أهداف عامة :

- مساعدة الأطباء غير ذوي الخبرة الكافية في تشخيص أورام الثدي.
- تحقيق المشاركة التقنية في الاستفادة من الخبرات البشرية الطبية النادرة.
- استخلاص الاستنتاجات من قواعد معينة ومن علاقات معقدة للتصرف.

2.5.1 أهداف خاصة:

- استخدام التقنيات الجديدة في تشخيص أورام الثدي
- توصيل المفاهيم الأساسية عن التنقيب عن البيانات .
- بناء نموذج يعمل علي تشخيص أمراض الثدي.

6.1حدود البحث :

تطبيق المشروعفي تشخيص أورام الثدي واكتشاف نوعية الورم باستخدام مجموعة بيانات تحتوي على قراءات مأخوذة من اجهزة الكشف عن أورام الثدي كما في الملاحق.

7.1 أهمية البحث:

تستفيد المراكز الصحية ومراكز تشخيص أورام الثدي و المستشفيات من هذا المشروع , مع جني عائد التسهيل و اليسر في تشخيص الأورام و زيادة الخبرة و المعرفة للكوادر الطبية العاملة في هذا المجال .

8.1 المساهمات المتوقعة من البحث:

من المتوقع أن يساهم البحث في حل بعض المشاكل التي ذكرت من قبل بالإضافة إلى ذلك فهويزيل الكثير من معاناة المرضى وعوائلهم بالإضافة لتوفير الجهد والوقت والتكاليف.

البـــــاب الثانـي

الخلفية النظرية و الدراسات السابقة

الفصل الأول:

1.2التنقيب عن البيانات:

1.1.2 تعريف تنقيب البيانات:

تستخدم عملية تنقيب البيانات أدوات مختلفة في تحليل البيانات لاكتشاف Patterns و العلاقات في بيانات قد تستخدم لوضع توقعات صحيحة.

أن الخطوة التحليلية الأولى في تنقيب البيانات هي وصف البيانات وتلخيص عناصرها الإحصائية (مثل الانحراف المعياري Standard Deviation) و المراجعة البصرية باستخدام اللوحات Charts و المخططات Graphs و البحث عن اتصالات ذات معنى بين المتغيرات Variables.

لكن وصف البيانات وحده لا يمكن أن يوفر خطة عمل , و انما يجب بناء نموذج توقعي Predictive Model مؤسس على بصمات Patterns محددة من نتائج معروفة , و بعد ذلك فحص ذلك النموذج على نتائج متحصلة من النموذج الأصلي . يجب أن لا يختلف النموذج الجيد أبدا عن الحقيقية(مثلا خارطة الطرق هي ليست تمثيل دقيق للطرق الحقيقية)لكن يمكن أن تكون موجه مفيد لفهم طبيعة العمل والخطوة الأخيرة هي إثبات Verifyالنموذج .[1]

يتضمن تنقيب البيانات استخدام أدوات دقيقة لتحليل البيانات لاكتشاف أشياء غير معروفة سابقاً وبصمات وعلاقات صحيحة في مجاميع بيانات كبيرة .و يمكن أن تتضمن هذه الأدوات نماذج إحصائية وخوارزميات رياضية لطرق التعلم بالحاسوب (خوارزميات تحسن أدائها بصورة أوتوماتيكية من خلال التجربة ، مثل الشبكات العصبية أو أشجار القرار) بالنتيجة فأن تنقيب البيانات هو أكثر من كونه تجميع و أدارة البيانات ، أنه يتضمن أيضاً تحليل Analysis وتوقع Prediction.[1]

يمكن أن يعمل تنقيب البيانات على بيانات ممثلة بشكل كمي أو نصي أو متعددة الوسائط.و تستطيع تطبيقات تنقيب البيانات أن تجري معاملات عديدة و متغيرة لفحص البيانات . تحتوي هذه التطبيقات على الارتباط Association ، تصنيف Classification ، و العنقدة Clusteringوالتوقع Forecasting .

يعتبر بعض المراقبين أن تنقيب البيانات هو خطوة في عملية أكبر تعرف باكتشاف المعرفة في قواعد البيانات Knowledge DiscoveryDataBase (KDD). [1]

و الخطوات الأخرى في عملية KDD بتسلسل تطوري تتضمن تنقية البيانات Cleaning ، تكامل البيانات Integration ، اختيار البيانات Selection ، نقل البيانات

Transformation ، تنقيب البيانات Data Mining ، تقييم البصمة Pattern Evaluation ، وتمثيل المعرفة Knowledge Representation .

وتتضمن نتائج تنقيب البيانات التنبؤ بما قد يحدث في المستقبل , و تصنيف الأشخاص أو الأشياء في مجاميع من خلال تمييز البصمات و اعتماداً على صفاتهم , كما تتضمن سلسلة ما هي الأحداث التي ستؤدي إلى أحداث قادمة .[1]

تقنية تنقيب البيانات هي ليستتحطيم كمية كبيرة من البيانات بشكل عشوائي , او تطبيق أعمى للخوارزميات , و لا يصعب فهمه او يحتاج إلي درجة علمية في علم الحاسوب , انما هي تقنية عملية تفاعلية والتي تستغل تقنيات التحليل والقدرة الحاسوبية لتجد علاقات لم يتم اكتشافها سابقاً من بصمات في البيانات , وهي تتطلب خبرة كبيرة.[1]

2.1.2أسس تنقيب البيانات:

أن تنقيب البيانات جاهز للتطبيق في مجتمع الأعمال لأنه مسند من قبل ثلاثة أسس :

- ❖ جمع بيانات كبيرة الحجم .
- ❖ حواسيب قوية ذات معالجات متعددة .
- ❖ خوارزميات تنقيب البيانات .[1]

3.1.2محددات تنقيب البيانات:

حتى تكون عملية تنقيب البيانات واضحة فأنها تحتاج إلي تقنية خبيرة ولأخصائي تحليل والذين يستطيعون هيكلة التحليل وترجمة الإخراج الذي تم توليده . بالنتيجة فأن محددات نجاح عملية تنقيب البيانات هي بصورة رئيسية البيانات أو الأفراد ، بدلاً من التقنية .

ومحدد آخر لتنقيب البيانات هو العلاقات فبينما يمكن تحديد الارتباطات بين السلوك و/أو المتغيرات ، فليس من الضروري تحديد علاقات عشوائية.[1]

4.1.2مجال تنقيب البيانات:

تستطيع تقنية تنقيب البيانات توليد فرص أعمال جديدة من خلال توفير هذه القدرات :

1. توقع أوتوماتيكي للأعمال والسلوك.
2. اكتشاف ممكنن لبصمات غير معروفة سابقاً .[1]

5.1.2التقنيات الأكثر استخداماً في تنقيب البيانات هي :

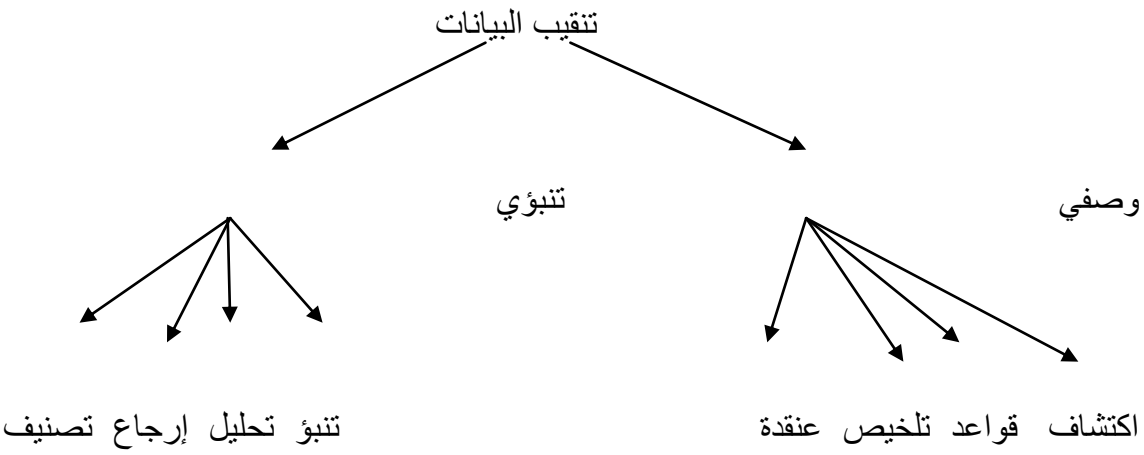
1. الشبكات العصبية الذكية Artificial Neural Networks .
2. أشجار القرار Decision Tree .
3. خوارزميات جينية Genetic Algorithms .
4. طريقة الجار الأقرب Nearest Neighbor .
5. حث القاعدة Rule Induction .

[1]

6.1.2 أهداف تنقيب البيانات :

أن أنواع انماطتنقيب البيانات التي يمكن اكتشافها تعتمد على أهداف تنقيب البيانات المستخدمة .

يوجد نوعين من أهداف تنقيب البيانات : أهداف تنقيب البيانات الوصفية التي تصف الخواص العامة للبيانات الموجودة و النوع الثاني أهداف تنقيب البيانات التنبؤية : التي تحاول عمل التنبؤات اعتمادا على البحث في البيانات المتوفرة .[1]



شكل(1.2) يوضح المعرفة المتنوعة في تنقيب البيانات

أنماط المعرفة في تنقيب البيانات :

1) التشخيص Characterization: تشخيص البيانات هو تلخيص للصفات العامة للمواضيع في الصنف الهدف ، وإنتاج مايسمى قواعد الشخصية .

2) التمييز Discrimination: تنتج بيانات التمييز Discrimination ما يسمى قواعد التمييز وهي أساساً المقارنة للصفات العامة للمواضيع بين صنفين يشار لها كصنف الهدف وصنف المفاضلة Contrasting.
3) تحليل الارتباط Association Analysis: تحليل الارتباطهو الاكتشاف لما يطلق عليه بصورة عامة قواعد الارتباط association rule. أنها تدرس تكرار العناصر سوية في قاعدة بيانات تعاملية ، تعتمد على سماح ما يسمى إسناد Support يحدد تكرار مجاميع العناصر.
4) التصنيف Classification: تحليل تنظيم البيانات في أصناف معطيات. يعرف أيضاً بالتصنيف المشرف عليه Supervised Classification يستخدم التصنيف مؤشرات الصنف المعطى لطلب المواضيع في مجموعة البيانات .
5) التنبؤ Prediction: لقد جلب التنبؤ الانتباه المثير حول حيث أعطى التنبؤ في محتوى العمل الناجح . يوجد نوعان عامان من التنبؤ :واحد يستطيع ما محولة تنبؤ بعض قيم البيانات غير المتوفرة والآخر يستطيع تنبؤ مؤشر صنف بعض البيانات.
6) العنقدة Clustering: مشابهة إلي التصنيف فهي تنظم البيانات في أصنافعلى كل حال ، بعكس التصنيف يكون في العنقدة مؤشرات الصنف هي غير معروفة وهي عائدة إلي خوارزمية العنقدة .
7) تحليل الدخيل Outlier analysis: الدخيل هي عناصر بيانات لا يمكن تجميعها في صنف معين أو عنقود محدد تسمى أيضاً كاستثناءات أو مفاجآت أنها غالبا ماتكون مهمة جدا بتحديدها .[1]

7.1.2 معمارية تنقيب البيانات:

تعمل العديد من أدوات تنقيب البيانات حالياً خارج منجم البيانات ، فأنها تحتاج الي خطوات أضافية لاستخلاص ، و جلب وتحليل البيانات , و يمكن استخدام منجم البيانات لتحسين عمليات التنقيب .[1]

8.1.2 تنقيب البيانات ومناجم البيانات:

دائما تستخلص البيانات المراد تنقيبها أولاً من منجم بيانات تابع الي وحدة معينة وتكون في تنقيب البيانات كقاعدة بيانات أو منبع بيانات .

الفصل الثاني :

2.2النظم الذكية :

من منتصف الخمسينات في القرن العشرين كان باحثو AI يقدمون الوعود ببناء آلات ذكية للأغراض العامة علي قاعدة معرفة تفوق ذكاء الإنسان

1.2.2تعريف المعرفة:

هي فهم نظري ، أو عملي لموضوع معين ، أو نطاق معين.كما ان المعرفة هي حاصل جمع ماهو معروف حاليا. يسمى هؤلاء الذين لديهم معرفة خبراء وهم الناس الاكثر قوة والاكثر اهمية لتنظيماتهم[2] .

2.2.2 تعريف الخبير بصفة عامة:

يمكن اعتبار أي فرد خبير في نطاق معين إذا كانت لديه معرفة عميقة وخبرة عملية قوية في نطاق معين,ويمكن أن يكون مجال النطاق محدودا.[2]

3.2.2 تفكير الخبراء:

تكون العملية الذهنية البشرية داخلية ومعقدة جدا في تمثيلها علي هيئة خوارزميات,الا ان معظم الخبراء يمكنهم التعبير عن معرفتهم في صورة قواعد لحل المشكلة.

يمكن صياغة القواعد الأساسية في صورة IF-THEN تسمي بقواعد الإنتاج.

يمكن تعريف مصطلح "قاعدة" في AI هي نوع من تمثيل المعرفة الأكثر استخداما وتوفر القاعدة بعض الوصف لكيفية حل المشكلة وتكون القواعد سهلة الإنتاج والفهم نسبيا.[2]

4.2.2 القواعد كأسلوب لتمثيل المعرفة:

تتكون أي قاعدة من جزأين: جزء IF ويسمي عنصرا شرطيا وجزء THEN ويسمي نتيجة منطقية ويأخذ التكوين الأساسي للقاعدة الشكل التالي:

IF <antecedent>

THEN <consequent>

5.2.2 اللاعبون الرئيسيون في فريق تطوير نظام الخبرة:

بمجرد أن يقدم الخبير البشري المعرفة يمكننا أن ندخلها الحاسب , ونتوقع من الحاسب أن يعمل كمساعد ذكي في نطاق خبرة معين، أو يحل المشكلة التي تحتاج في حلها إلي خبير ونريد أن يكون الحاسب قادرا علي عمل تكامل للمعرفة الجديدة، ويبين معرفته في صورة سهلة القراءة والفهم ويتعامل مع جمل بسيطة بلغة طبيعية بدلا من لغة البرمجة الصناعية.

أخيراً نريد من الحاسب أن يشرح كيف وصل إلي استنتاج معين , وبكلمات أخري علينا أن نبني نظام خبرة يمكنه العمل علي مستوي الخبير البشري في مجال نطاق ضيق.[2]

1.5.2.2 أعضاء فريق تطوير نظام الخبرة:

1. خبير النطاق.
2. مهندس المعرفة.
3. المبرمج.
4. مدير المشروع.
5. المستخدم النهائي.[2]

2.5.2.2أجزاء نظام الخبرة:

للنظام الخبرة المبني علي القواعد خمسة أجزاء:

1. قاعدة المعرفة.
2. قاعدة البيانات.
3. آلة الاستدلال.
4. وسيلة التوضيح.
5. الشكل البيني للمستخدم.[2]

6.2.2 الخواص الأساسية لنظام الخبرة:

يبنى نظام الخبرة ليعمل علي نفس مستوي الخبير البشري في نطاق متخصص لذلك الخاصية الأكثر أهمية لنظام الخبرة هي أداؤه المرتفع الجودة. إذ أن المستخدم لن يرضي بالنظام إذا كانت النتائج خطأ.

السمة الفردية لنظام الخبرة هي إمكانية التوضيح explanation capability وهذه تجعل نظام الخبرة قادراً على مراجعة تفكيره الخاص به ، وتوضيح قراراته . [2]

1.6.2.2احتمالية خطأ النظم الخبيرة:

حتى الخبير البارز ليس إلا إنسانا ، وبالتالي يمكن أن يقع في الخطأ . ويقترح في نظام الخبرة الذي ينبنى لأداء العمل بمستوى الخبير البشري يجب أن يسمح له بالوقوع في بعض الأخطاء أيضا .إلا أننا نظل نثق في الخبراء ،حتى مع معرفتنا أن أحكامهم تخطئ في بعض الأحيان. وبالمثل على الإقل في أغلب الحالات يمكننا أن نعتمد على الحلولالتي توفرها نظم الخبرة مع إمكانية حدوث أخطاء ، ويجب أن نكون ملمين بذلك .[2]

7.2.2 أساليب استدلال التسلسل للأمام والتسلسل للخلف:

في نظام الخبرة المبني على القواعد تمثل معرفة النطاق بمجموعة من قواعد إنتاج IF-THEN وتمثل البيانات بمجموعة من الحقائق عن الحالة الحالية , وتقارن آلة الاستدلال كل قاعدة مخزنة في قاعدة المعرفة مع الحقائق المضمنة في قاعدة البيانات وعندما يتفق جزء IF(condition) من القاعدة مع الحقيقة فإنه ينفذ جزء THEN الخاص بها ,ويمكن أن نغير القاعدة التي أطلقت مجموعة من الحقائق عن طريق إضافة حقيقة مثلاً .[2]

1.7.2.2التسلسل للأمام:

يكون التسلسل للأمام تفكيرا تديره البيانات . يبدأ التفكير من بيانات معروفة ويستمر للأمام وفي كل مرحلة تنفذ القاعدة الموجودة في القمة فقط , وعند إطلاق القاعدة فإننا نضيف حقيقة جديدة في قاعدة البيانات . ويمكن تنفيذ أي قاعدة مرة واحدة فقط وتتوقف دورة التوفيق والإطلاق عندما لا يصبح من الممكن إطلاق قواعد أخرى .[2]

2.7.2.2 التسلسل للخلف:

يكون التسلسل للخلف تفكيرا يديره الهدف ويكون لنظام الخبرة في التسلسل للخلف هدف (حل فتراضي) وتحاول آلة الاستدلال أن تجد دليلا لتثبته , وفي البداية يتم البحث في قاعدة المعرفة لإيجاد القواعد التي يمكن أن لها الحل المطلوب .[2]

الفصل الثالث:

3.2النظم الهجين:

1.3.2 تعريف النظام الهجين:

النظام الذكي الهجين هو نظام يدمج تقنيتين على الأقل . مثال لذلكينتج عن دمج الشبكة العصبية مع النظام الضبابي نظام عصبي ضبابي .

ويمثل دمج التفكير الاحتمالي ، والمنطق الضبابي ، والشبكات العصبية ، والحوسبة الارتقائية قلب الحوسبة اللينة (sc) soft computing وهي طريقة منبثقة لبناء النظم الذكية الهجين القادرة على التفكير ، والتعلم في بيئة غير دقيقة ، وغير مؤكدة .

وكان لطفي زادية LotfiZadeh "أب المنطق الضبابي" أول من تمكن من إمكانيات الحوسبة اللينة.[2]

2.3.2الحوسبة اللينة:

بينما تستخدم الحوسبة التقليدية أو "الصلبة" قيما واضحة أو أعداد ، فان الحوسبة اللينة تستطيع العمل مع معلومات غير مؤكدة ، وغير دقيقة ،وغير كاملة بطريقة تعكس التفكير البشري.[2]

3.3.2 الفرق بين الحوسبة اللينة والذكاء الصناعي:

يحاول الذكاء الصناعي التقليدي التعبير عن معرفة البشر بمصطلحات رمزية , ويعد نظام الخبرة المنتج الأكثر نجاحاً للذكاء الصناعي التقليدي . إلا أن نظام الخبرة يكون جيدا إذا أمكن الحصول على المعرفة.

إلا انه خلال البضع سنوات الماضية اتسع نطاق الذكاء الصناعي بسرعة ليشمل الشبكات العصبية الصناعية ،والخوارزميات الجينية ،وحتى نظرية الفئات الضبابية. وهذا جعل الحدود بين الذكاء الصناعي الحديث والحوسبة اللينة مبهمة ومضللة .[2]

4.3.2 مكونات النظام الخبير الهجيني:

لابد أن يكون النظام الهجين الجديد جيداً , ولذلك من الضروري اختيار المكونات الصحيحة لبناء النظام الهجين الجديد .

لكل مكون نقاط قوة وضعفه , ويهتم التفكير الاحتمالي أساسا بعدم التأكد ويهتم المنطق الضبابي بعدم الدقة ، والشبكات العصبية بالتعلم ، والأمثلة الارتقائية بمجموعة الامثال . ويدمج النظام الهجين الجيد مميزات هذه التقنيات مع بعضها البعض , ويسمح تعاونها للنظام الهجين باحتوائه على الإحساس العام ، واستخلاص المعرفة من البيانات الخام ، واستخدام آليات تفكير شبيهة بالإنسان ، والتعامل مع عدم التأكد وعدم الدقة ، وتعلم التكيف مع التغييرات السريعة والبيئة غير المعروفة [2].

5.3.2نظم الخبرة العصبية:

تشترك نظم الخبرة والشبكات العصبية كتقنيات ذكاء في الأهداف العامة . فيحاول كل منهما تقليد الذكاء البشري، وربما إنتاج آلة ذكية. إلا أنهما يستخدمان وسائل مختلفة جدا لتحقيق أهدافهما فبينما تعتمد نظم الخبرة على الاستدلالات المنطقية وأشجار القرار وتركز على تمثيل التفكير البشري فإن الشبكات العصبية تعتمد على التشغيل المتوازي للبيانات وتركز على تمثيل مخ الإنسان. وتعامل نظم الخبرة المخ كصندوق أسود بينما تنظر الشبكات العصبية إلى هيكله ووظائفه وبصفة خاصة إلى مقدرته على التعلم. وتنعكس هذه الاختلافات الأساسية في أساليب تمثيل المعرفة وتشغيل البيانات المستخدمة في نظم الخبرة والشبكات العصبية.

وتمثيل المعرفة في نظم الخبرة المبنية على القواعد بواسطة قواعد إنتاج IF - THEN التي يتم جمعها عن طريق مشاهدة الخبراء البشريين، أو عقد لقاءات خاصة معهم. إضافة إلى ذلك بعد تخزين القواعد في قاعدة المعرفة لا يمكن تعديلها حتى بواسطة نظام الخبرة نفسه. ولا يمكن أن تتعلم نظم الخبرة من خلال الخبرة، أو تتكيف للبيئات الجديدة. والإنسان البشري فقط هو القادر على تعديل قاعدة المعرفة عن طريق إضافة بعض القواعد، أو التغيير فيها، أو حذفها.[2]

6.3.2 التفكير التقريبي:

في نظام الخبرة المبني على القواعد تقارن آلة الاستدلال جزء الشرط من كل قاعدة مع البيانات المقدمة في قاعدة البيانات . وعندما يتفق جزء IF للقاعدة مع البيانات الموجودة في قاعدة البيانات تنطلق القاعدة ، وينفذ جزء THENالخاص بها وفي نظم الخبرة المبنية على القاعدة ، يلزم اتفاق دقيق precise ونتيجة لذلك لا يمكن أن تعمل آلة الاستدلال مع بيانات غير كاملة أو مشوشة.[2]

الفصل الرابع :

4.2الدراسات السابقة:

1.4.2مرشد طبي خبير متعدد الوسائط:

يتكون النظام من مجموعة من الوحدات التي تتعاون فيما بينها لتقديم الحلول الطبية المناسبة لأمراض الحج وذلك بعد ان يتم تحميل النظام الخبير وتدريبه بالمعلومات والأسس والقواعد الضرورية اللازمة التي تعين النظام الخبير على تحديد نوع المرض. ثم يقدم النظام بعد ذلك الأرشادات العلاجية الموثوق فيها،ويقترح العلاج المناسب للحالة المرضية التي يتم استشارته في علاجها. ومن ضمن تلك البيانات التي يتطلبها النظام ؛ على سبيل المثال ؛ الحالة المرضية لتاريخ المريض ودرجة الحرارة ، ومقياس الضغط ، ونسبة السكر في الدم ، وبعض اعراض المرض .ومن ثم يعرض النظام على شاشة العرض المرئية بعض الشرائح التوضيحية او فيلم فيديو يبين الحالة المرضية ويشخصها ثم يقترح العلاج اللازم . وقد تم فحص وتقييم اداء النظام بعد بنائه من قبل بعض الأساتذة المتخصصين في امراض الحج لقياس مدى نجاح النظام الخبير في مجال التدريب والتعليم في مجال التدريب والتعليم في مجال امراض الحج .

2.4.2 استخدام تقنية التنقيب في البيانات لمرضي القلب:

كل سنة يموت الآلاف من الناس بسبب التوقف المفاجئ للقلب حسب تقرير منظمة الصحة العالمية، لهذا جاء البحث لتحسين التشخيص الطبي للأمراض ،والحد من الوقت للضغط علي الأطباء.لذلك بدأت أغلب المؤسسات الطبية بلستعمال أدوات التنقيب عن البيانات.

يقوم النظام بعمل توصيل المفاهيم الأساسية عن التنقيب عن البيانات ، مساعدة الأطباء غير ذوي الخبرة الكافية في تشخيص بعض الأمراض الخطيرة علي حياة الإنسان مثل أمراض القلب ، استخدام أنظمة ذكية و استخدام التقنيات الجديدة في تشخيص مرض القلب ، الاستفادة من المعلومات المخفية في قاعدة البيانات و المرتبطة بأمراض القلب ، بناء نموذج يعمل علي تشخيص أمراض القلب.

البـــــاب الثالث

الأدوات و تحليل النظام

الفصل الأول:

1.3مقدمة:

نتناول في هذا الباب الادوات والتقنيات المستخدمة وتحليل النظام .

2.3الأدوات المستخدمة

1.2.3 Microsoft EXCEL:

الأكسل EXCELهو احد البرامج التي أنتجتها شركة مايكروسوفت العملاقةهو عبارة عن برنامج الجداول الإلكترونية الحسابية والرياضية.يساعد المستخدم على بناء الجداول بكل سهولة وإجراء العمليات الحسابية البسيطة والمعقدة بكل سهولة ويسر,وتستخدم عاده لتخزين البيانات الحسابية .

2.2.3 Clementine:

هو نموذج يستقل بيانات مختلفة ويقوم بتحليلات متنوعة تساعد على بناء نماذج التنبؤ الدقيقة بسرعة وبشكل حدسي ، دون الحاجة إلي البرمجة وتصور بسهولة عملية استخراج البيانات وذلك باستخدام واجهة رسومية بديهية .
من هذه الواجهة يمكن الوصول بسهولة لبيانات منظمة وغير منظمة من مجموعة متنوعة من المصادر ، مثل قواعد البيانات التشغيلية واستخدام النمذجة لتحسين نتائج الأعمال في مجالات متنوعة في إدارة علاقات العملاء والتسويق وتخطيط الموارد و البحث الطبي والقانون والأمن القومي .
إنتاج تنبؤات أعمق وأكثر دقة من خلال الاستفادة من جميع البيانات لخلق صورة كاملة من المكونات و الاختيار بين مجموعة كاملة من الوظائف التحليلية المتقدمة ، بما في ذلك الفن والخوارزميات وإعداد البيانات الآلي ، وقدرات التصور التفاعلية الغنية .

ميزات رئيسية :

1. الأفكار الرئيسية واضحة ، واستخدامها لحل مشاكل الأعمال الحقيقية.
2. يقوم النموذج بحل أصعب التحديات .
3. له واجهة رسومية واضحة وسهلة.
4. يوفر الوقت في الحصول على نتائج أسرع من خلال أتمتة البيانات.
5. استخدام جميع البيانات.[5]

3.2.3 ExsysCorvid:

هو أداة قوية ، ثبت على نطاق واسع أنها مفيدة في بناء تطبيقات النظم الخبيرة التفاعلية على الانترنت. وهي مصممة لتكون سهلة التعلم لغير المبرمجين . أنها تستخدم منطقصنع القرار ومحرك الاستدلال لتحويل النظام لشكل منظم .

تستخدم في بناء الأنظمة الصغيرة والكبيرة أيضاً بسبب سهولة استخدامها.يمكن ان تستغرق بضع ساعات لبناء الأنظمة الكبيرة.

Corvidيحل المشاكل التالية:

1. يستخدم منطق صنع القرار.
2. التكامل مع تكنولوجيا المعلومات الأخرى.
3. واجهة المستخدم تمتاز بالقوة.
4. يوفر طرق متعددة لوصف المنطق.

قد يكون للنظام المعقد العديد من القواعد ، وهناك العديد من الطرق لبناء منطق للنظام .وهذا يسمح للقواعد التي تنظم الطريقة التي يبنى بها النظام الخبير حول نطاق المشكلة ويتم ربطها بأحدى طرق البرمجة مثل (java or html).

دمج تكنولوجيا المعلومات مع الموارد الأخرى :
غالبا ما تحتاج النظم الخبيرة للتكامل مع الموارد الأخرى للحصول على بيانات، أوحفظ نتائج عمليات. Corvid يجعل من السهل دمج تكنولوجيا المعلومات مع الموارد الأخرى إضافة و أي وظائف خاصة أخرى .

Corvid يمكن أن يبنى بسرعة.بطريقة بسيطة من قواعد تمكن خبراء المجال بناء نظم للعمل بكفاءة مع فرق التطوير .يمكن إضافة واجهات المستخدم لتتناسب مع موقع على شبكة الانترنت و إضافتها من قبل فريق التصميم و تكون مستقلة عن منطق النظام و تمتازهذه الواجهات بالمرونة وسهولة الإعداد.

4.2.3 HTML:

هي لغة بسيطة لبناء صفحات الإنترنت ، يعني ذلك أن أغلب الصفحات التي تزورها على شبكة الإنترنت تعتمد على هذه اللغة وقد يتخللها بعض السطور البرمجية للغات أخرى و HTML هي اختصار لـHyper Text Markup Languageوهي لغة سهلة جداً.

الفصل الثاني:

3.3 التحليل:

لدنيا مجموعة بيانات مأخوذة من قراءات 9 اجهزة لفحص اعراض اورام الثدي , لمجموع 104 مريضة , وتم وضع هذه البيانات في جدول EXCEL حيث تمثل الصفوف عدد المرضى , بينما تمثل الاعمدة قراءات اجهزة الكشف .

1.3.3 استخدام الكلمنتاين:

أفضل 10 خوارزميات في Data Mining وفقا ل IEEE (ICDM) الذي أقيم في ديسمبر 2006 في هونج كونج :

- C4.5
- K-Means
- SVM
- Apriori
- EM
- PageRank
- AdaBoost
- KNN
- NaiveBayes
- CART

والخوارزمية المستخدمة في هذا البحث هي خوارزمية C4.5 .[4].

1.1.3.3 خوارزمية C4.5 :

1.1.1.3.3مقدمة :

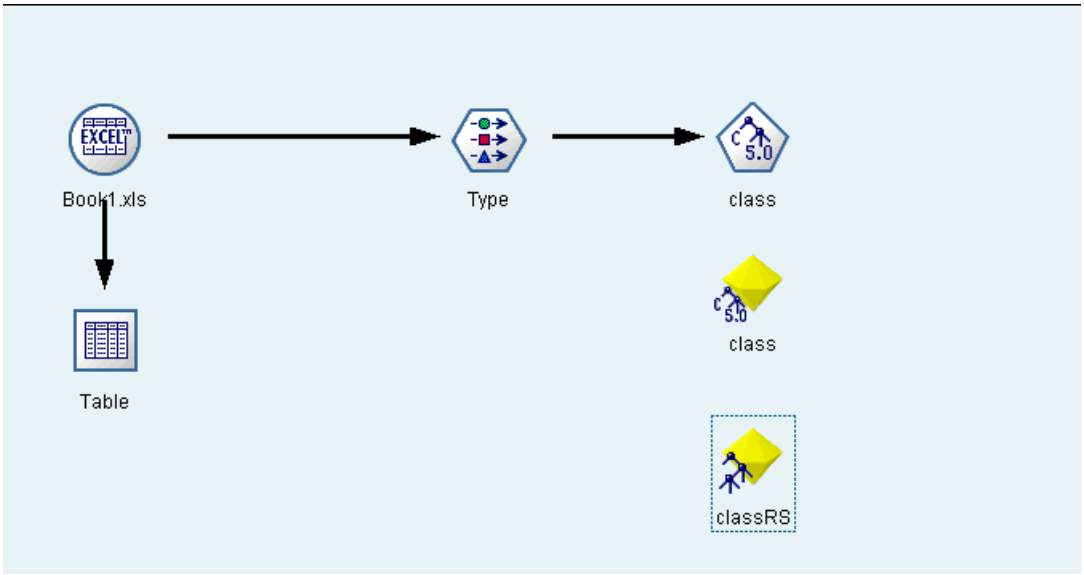
هي احد أنواع الخوارزميات التي تعمل بأسلوب شجرة القرار , حيث تتمثل فيها بتمثيل مجاميع القرارات وهذه القرارات تولد فوانين التصنيف الخاصة بمجاميع البيانات , ولشكل آخر فإنها تعمل غلى تصنيف الحالات , و هي من طرائق التصنيف شائعة الاستخدام . و هي تعمل بأسلوب (فرق-تسد) , إذ أنها تجزئ المسالة المعقدة إلى مسائل ابسط , ثم يتم الاستدعاء الذاتي للدالة نفسها ولكل الأجزاء , وبجمع حلول المسائل المجزاة

يتم الحصول على حل المسالة المعقدة , ومميزاته ا انها تتعامل مع الأرقام، الخصائص،الأرقام المفقودة والبيانات المشوشة .[4]

See5/C5.0 2.1.1.2.2:

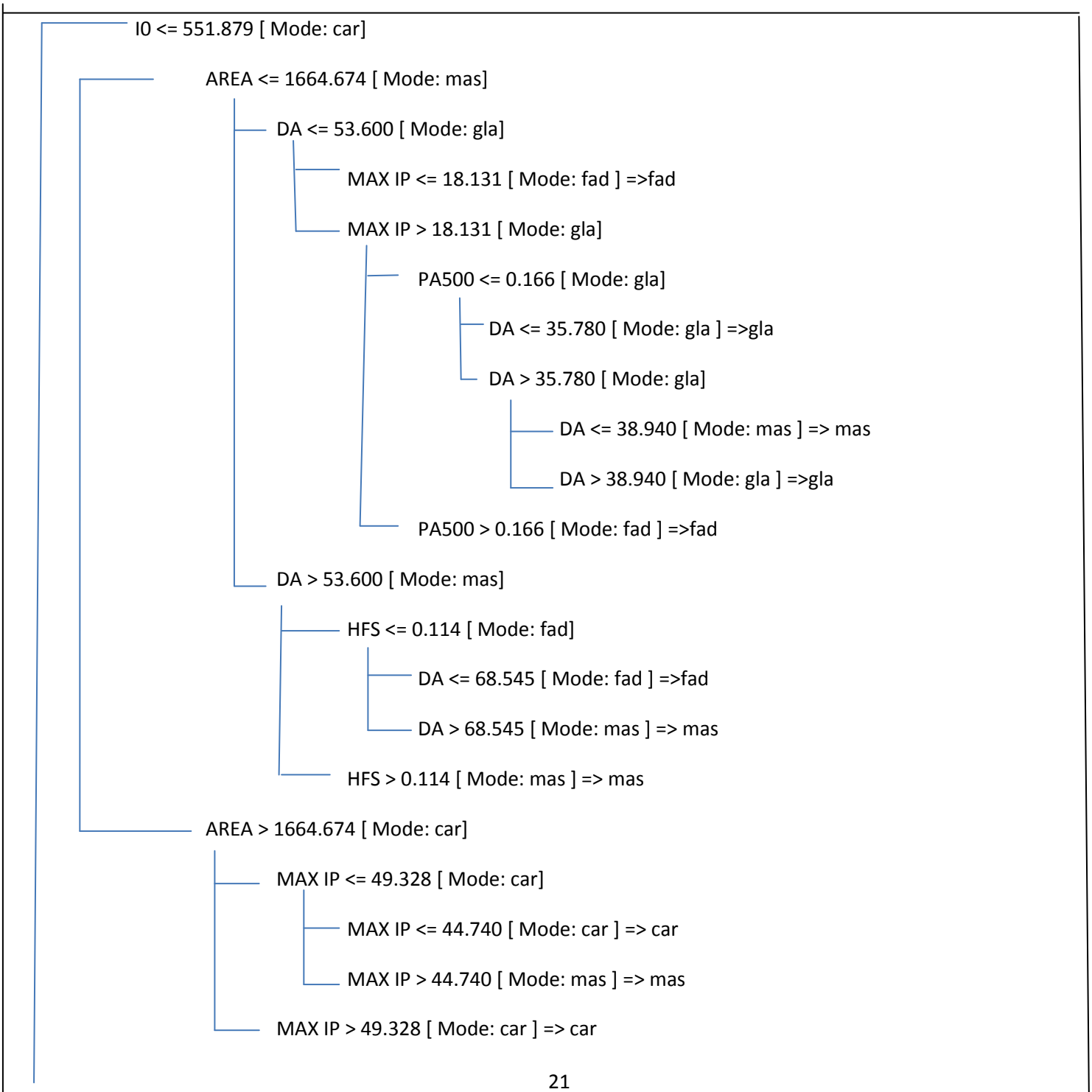
حلت محل C4.5 في عام 1997 و التغيرات التي حدثت فيها تشمل قابلية جديدة أدت لتحسن كفاءتها ، مثل :

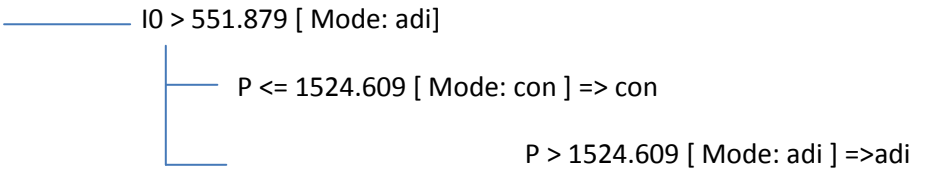
- أنواع بيانات جديدة (مثلا التواريخ) .
- القواعد الغير مرتبه
- تحسنت كثيرا قابلية كل من أشجارالقرار و rulesets .[4]



شكل(1.3) يوضح استخدام الكلمنتاين

القواعد


I0 <= 551.879 [Mode: car]
AREA <= 1664.674 [Mode: mas]
DA <= 53.600 [Mode: gla]
MAX IP <= 18.131 [Mode: fad] =>fad
MAX IP > 18.131 [Mode: gla]
PA500 <= 0.166 [Mode: gla]
DA <= 35.780 [Mode: gla] =>gla
DA > 35.780 [Mode: gla]
DA <= 38.940 [Mode: mas] => mas
DA > 38.940 [Mode: gla] =>gla
PA500 > 0.166 [Mode: fad] =>fad
DA > 53.600 [Mode: mas]
HFS <= 0.114 [Mode: fad]
DA <= 68.545 [Mode: fad] =>fad
DA > 68.545 [Mode: mas] => mas
HFS > 0.114 [Mode: mas] => mas
AREA > 1664.674 [Mode: car]
MAX IP <= 49.328 [Mode: car]
MAX IP <= 44.740 [Mode: car] => car
MAX IP > 44.740 [Mode: mas] => mas
MAX IP > 49.328 [Mode: car] => car

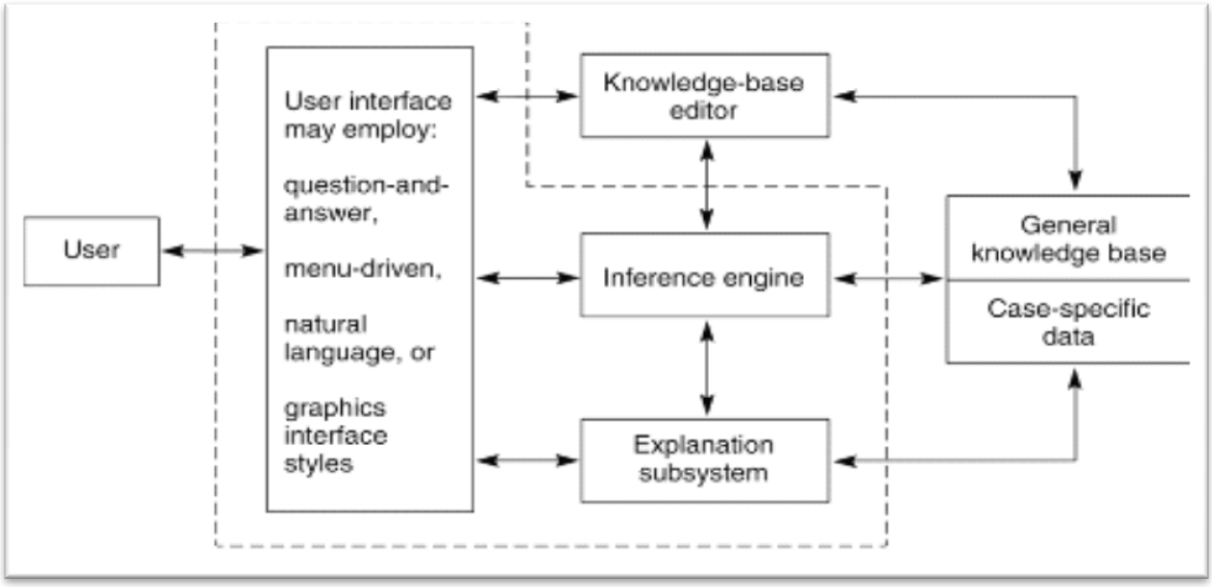


شكل رقم(2.3) يوضح القواعد الناتجة من خوازمية C5

شكل(3.3) يوضح تحليل القواعد

2.3.3 بناء النظم الخبيرة :

لكي يكون المستخدم قادر على التعامل مع النظام يجب أن يحتوي النظام الخبير على واجهه يستطيع معها كلا من المستخدم والنظام على الحديث والحوار مع أواسهلها استخداما هي سؤال وجواب (question- and- answer)حيث يسأل النظام والمستخدم يجيب . وهذه خطة العمل الذي ينبغي أن نتبعها في هذا المشروع .



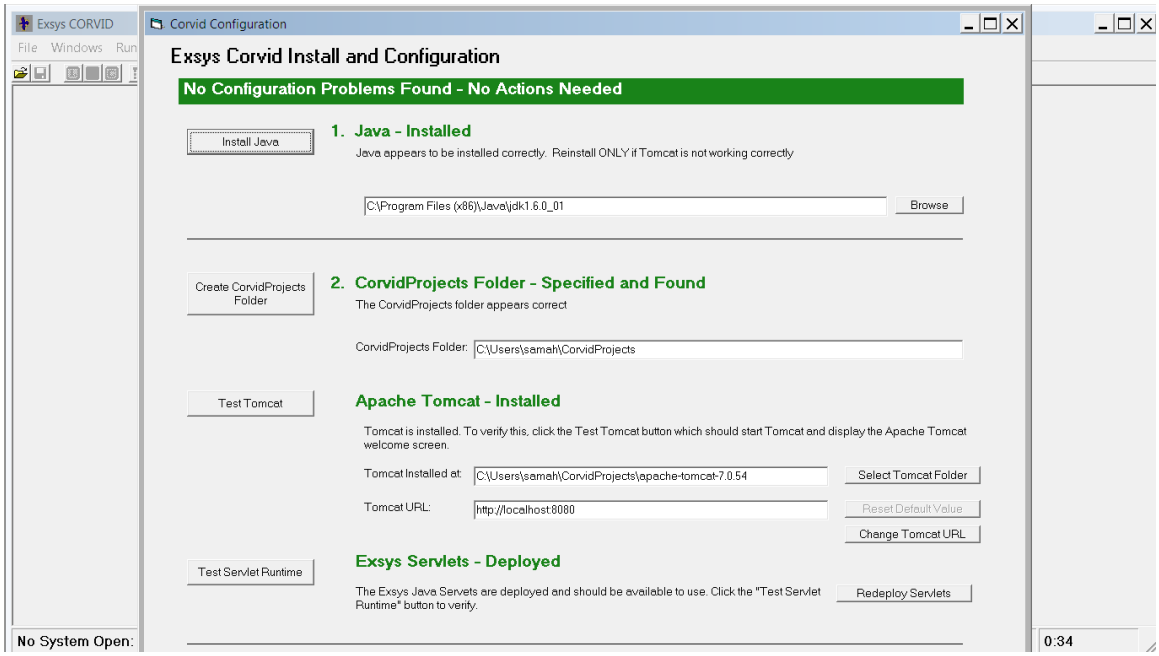
شكل (4.3) يوضح interface لنظام خبير

طريقة بناء النظام الخبير في ExsysCorvid :

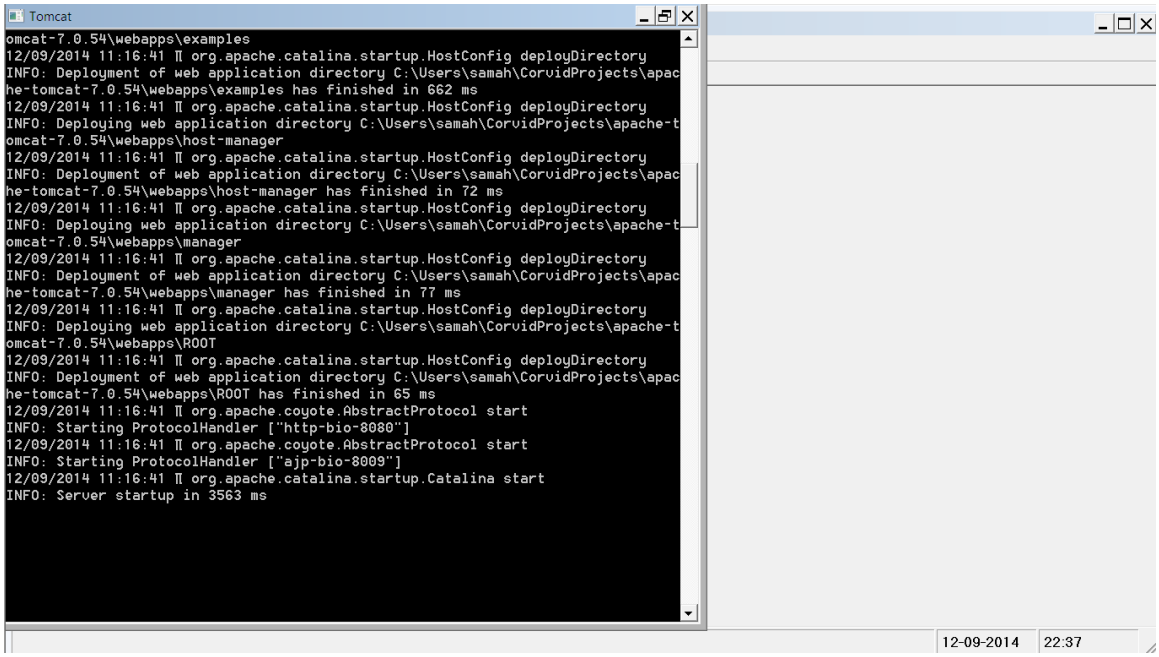
اولا نقوم بفتح ExsysCorvid

شكل رقم (5.3) يوضح مقدمة ExsysCorvid

شكل رقم (6.3) يوضح شاشة الصلاحية فيExsysCorvid

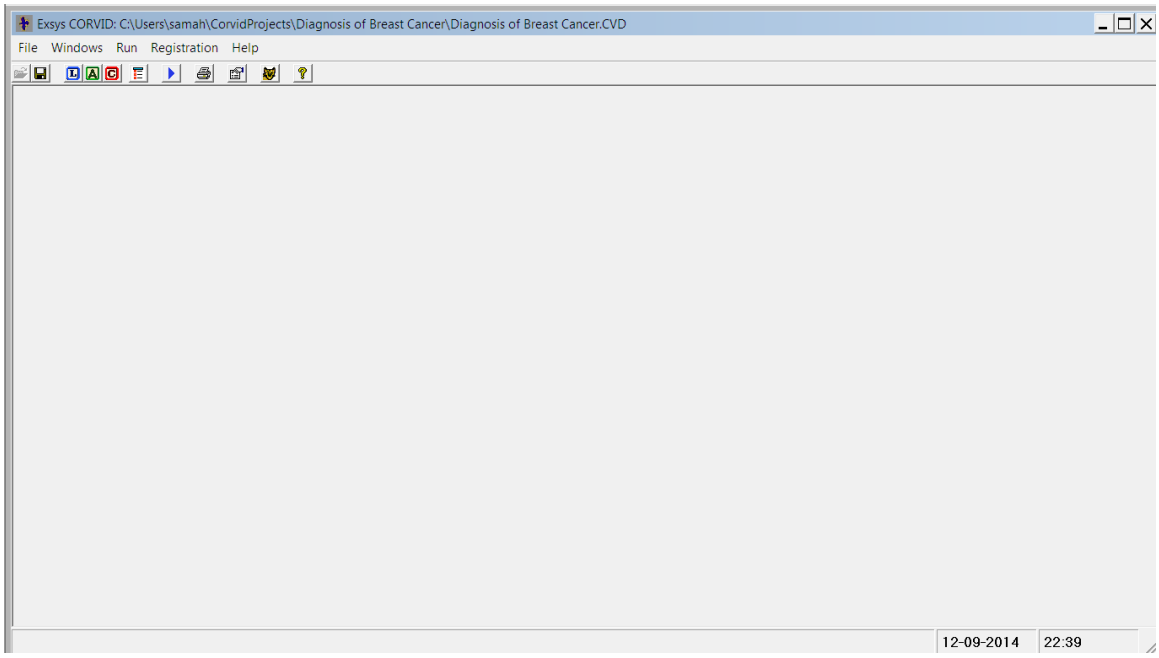


شكل رقم (7.3) يوضح ضبط ExsysCorvid



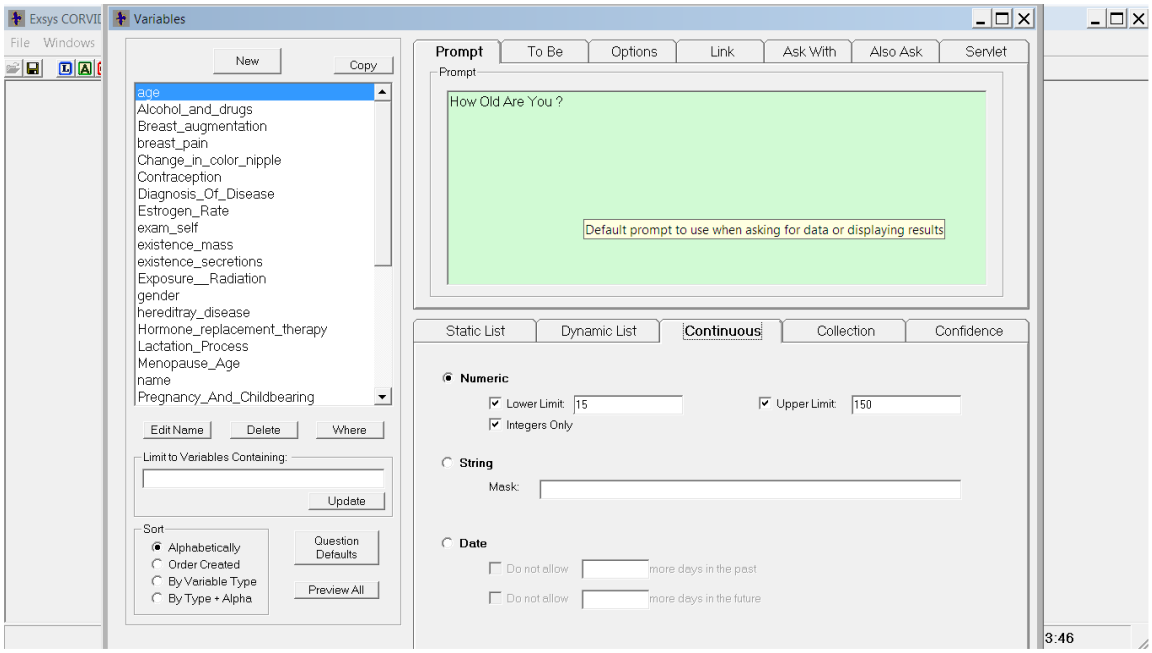
شكل رقم (8.3) يوضح ضبط TomCat

وبعد هذا تصبح الاداة جاهزة للاستخدام

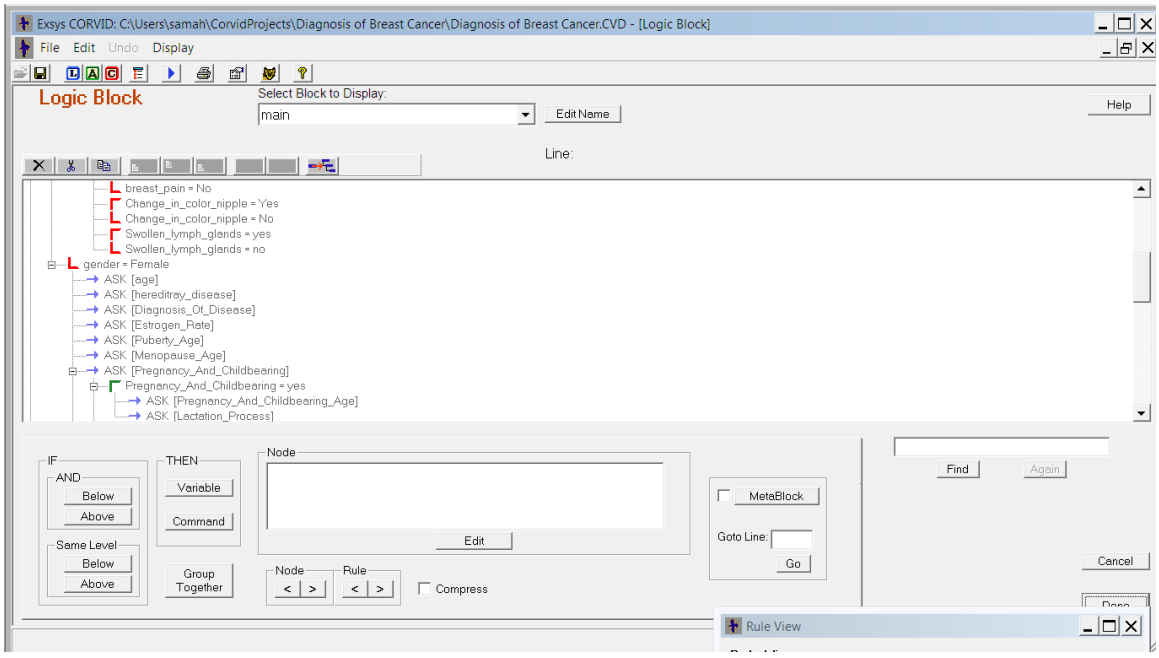


شكل رقم (9.3) يوضح الشاشة الرئيسية في ExsysCorvid

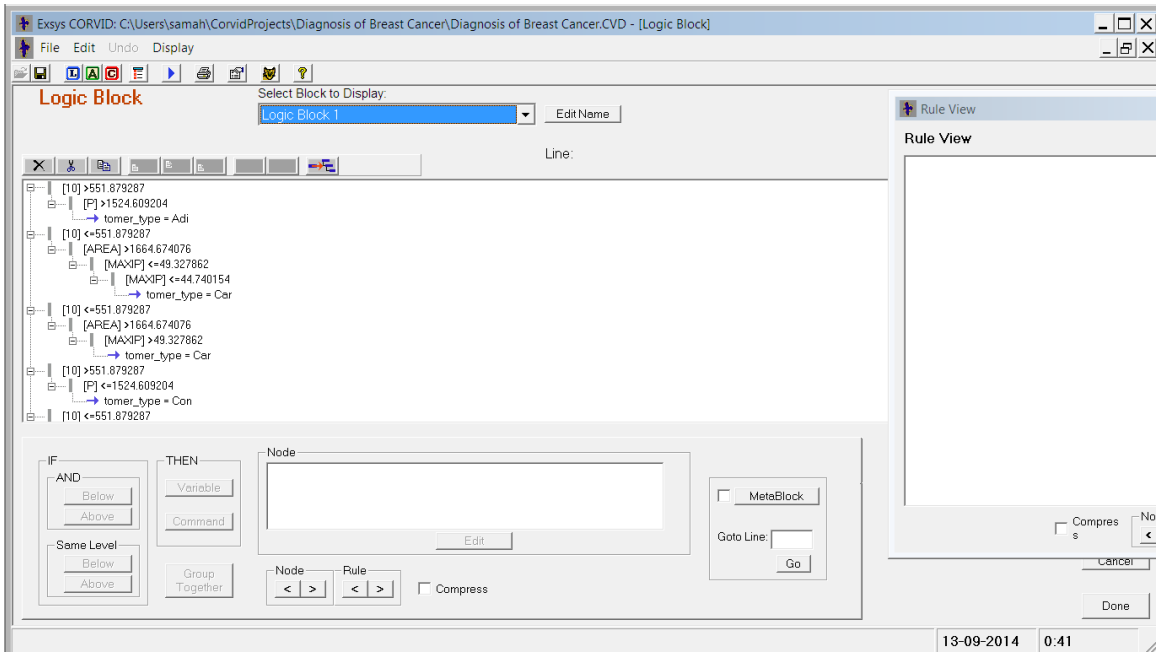
و في الخطوة الاولى لبناء النظام قمنا بتعريف المتغيرات



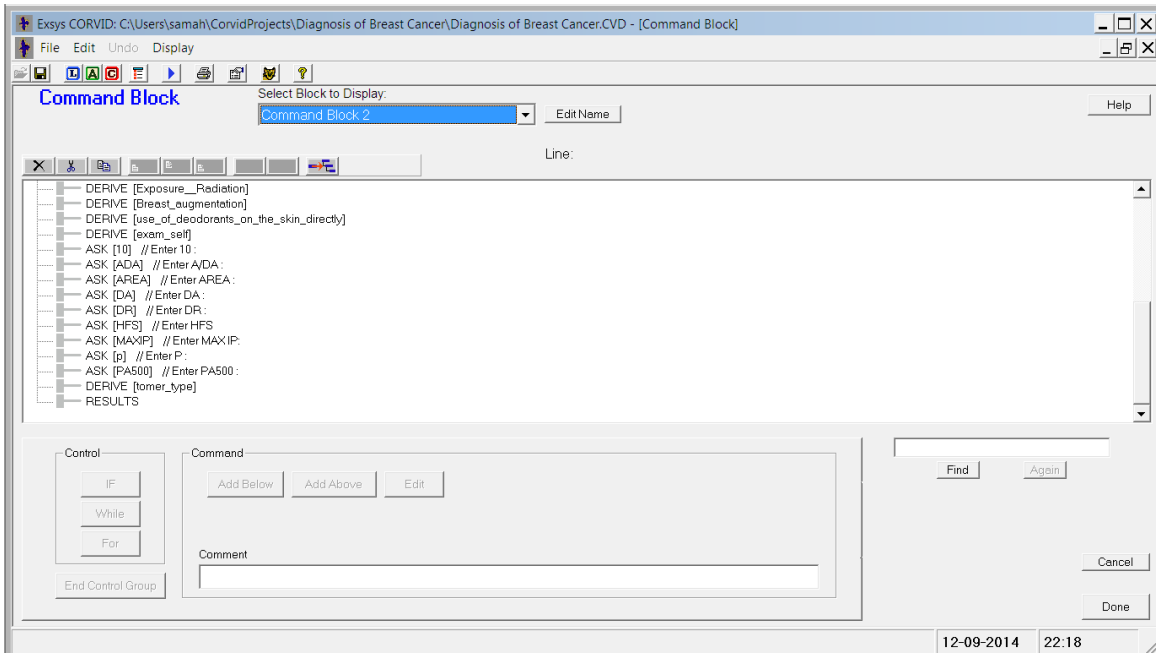
شكل رقم (10.3) يوضح إنشاء المتغيرات



شكل رقم (11.3) يوضح إنشاء الاوامر المنطقية للمتغيرات



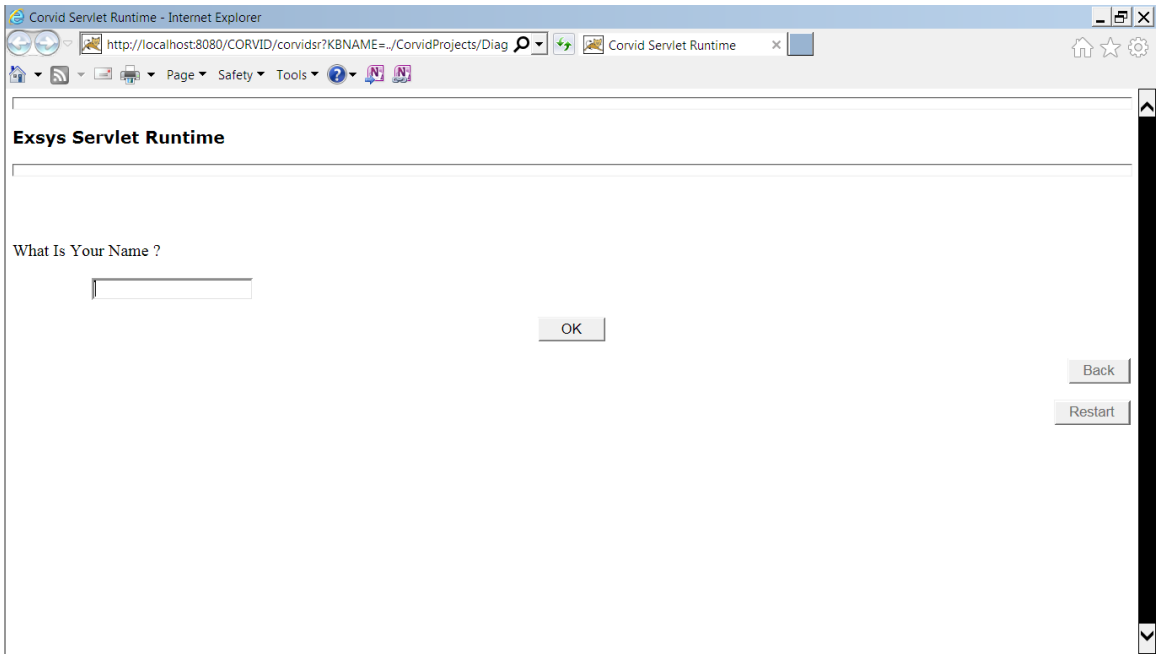
شكل رقم (12.3) يوضح إنشاء الاوامر المنطقية للقواعد



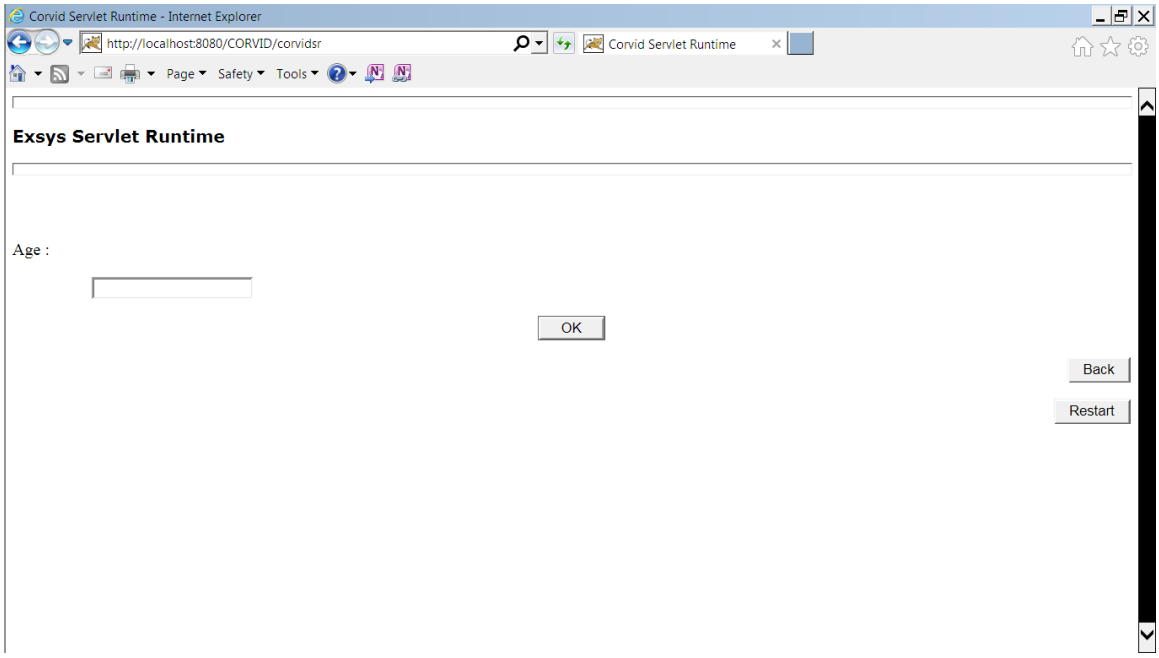
شكل رقم (13.3) يوضح إنشاء اوامر الترتيب

و يتم انشاء الواجهات التفاعلية بلغة HTMLوربطها مع اداة البناء .

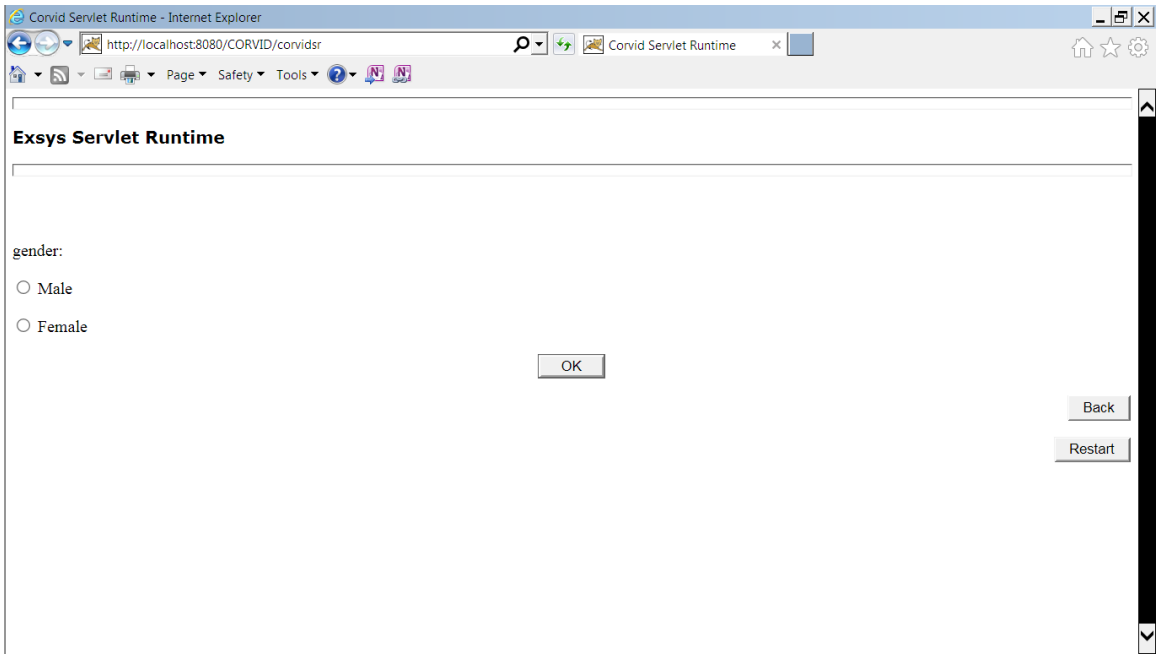
في مرحلة التنفيذ يحدث الاتي :



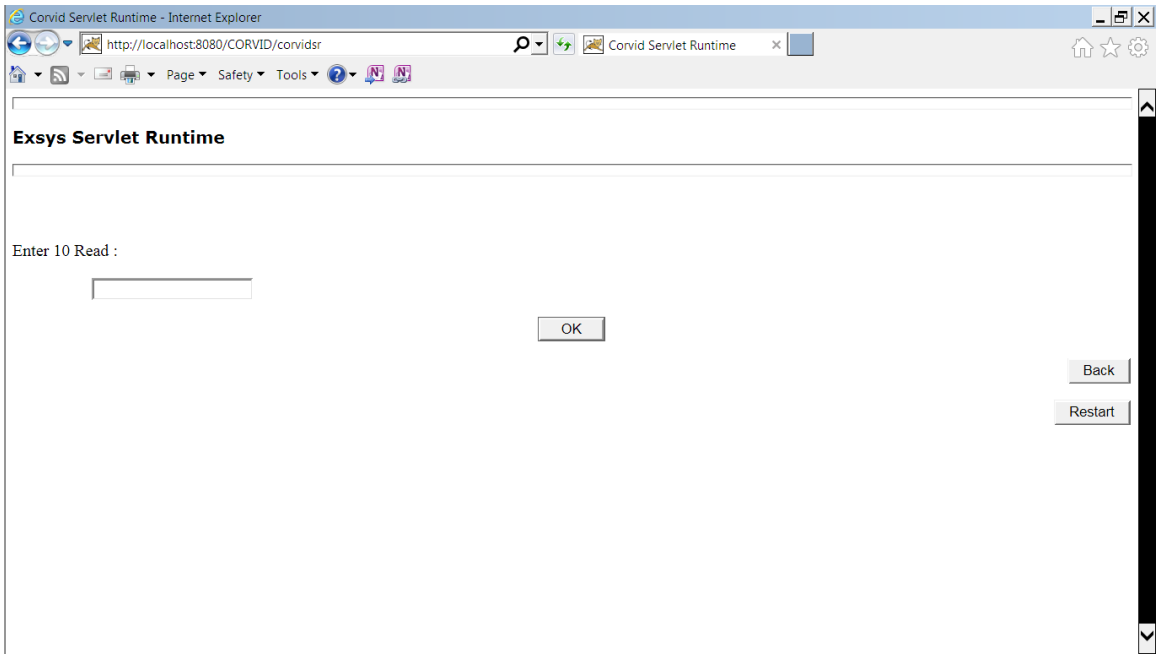
شكل رقم(14.3) يوضح ادخال اسم المريض



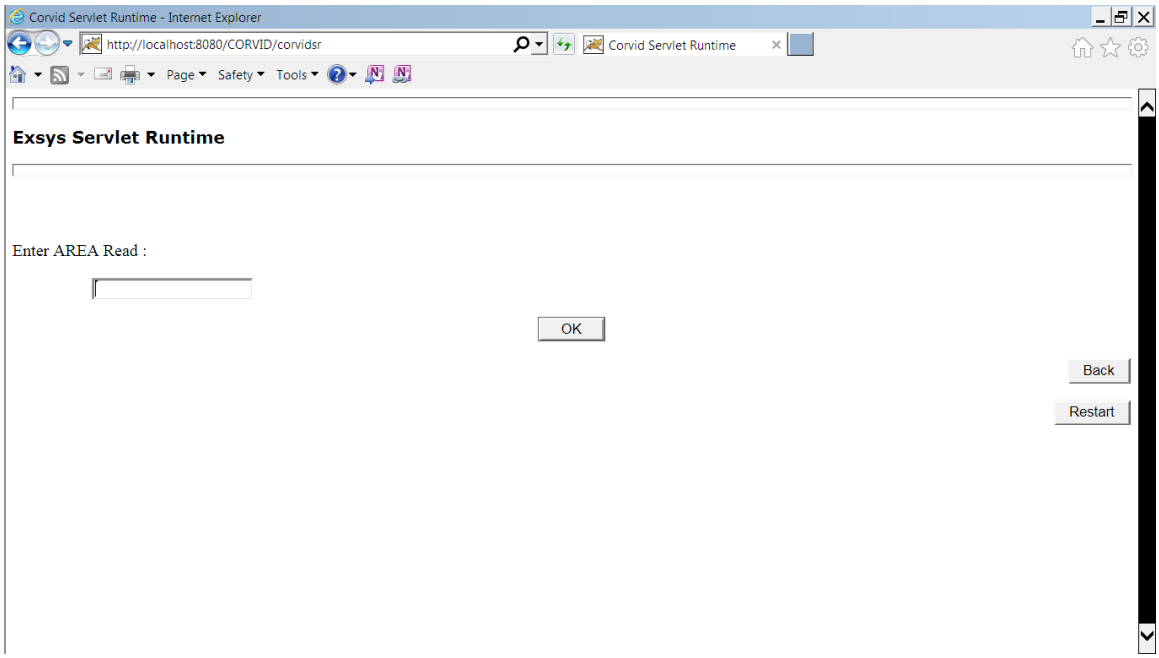
شكل رقم(15.3) يوضح ادخال عمر المريض



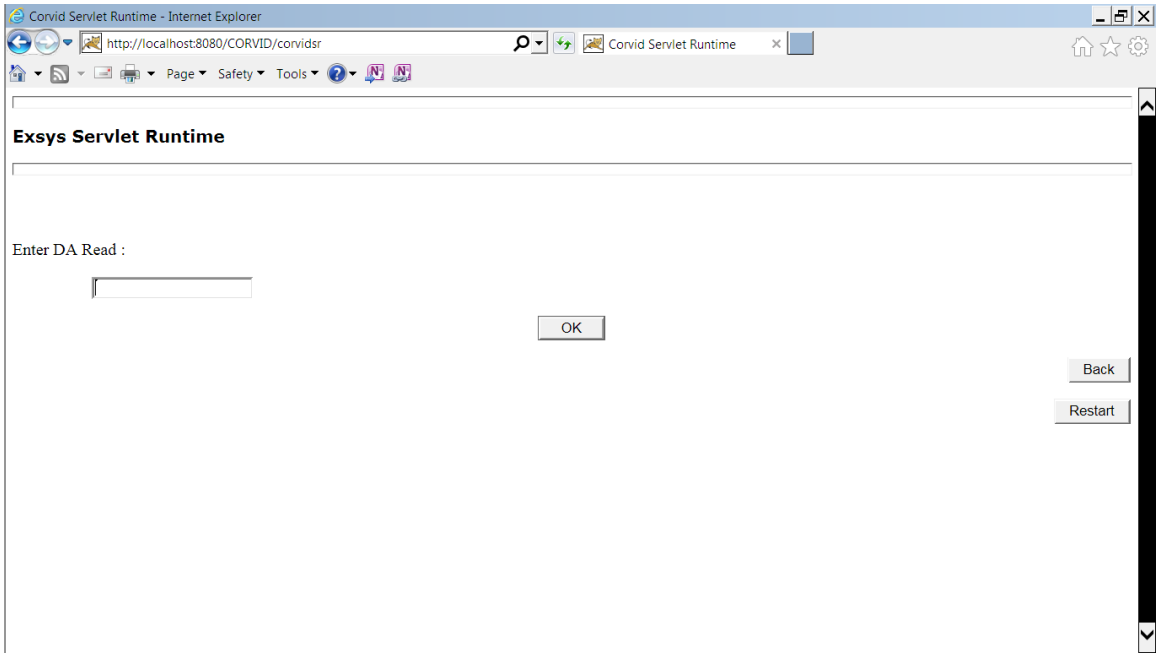
شكل رقم(16.3) يوضح تحديد جنس المريض



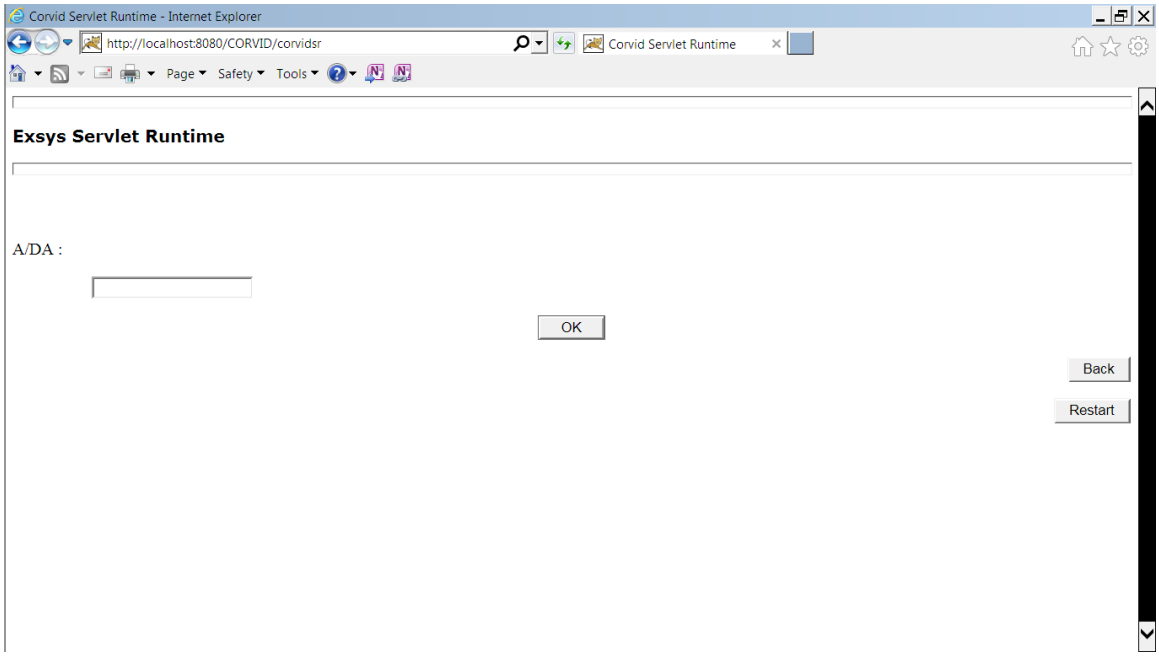
شكل رقم(17.3) يوضح ادخال قراءة 10



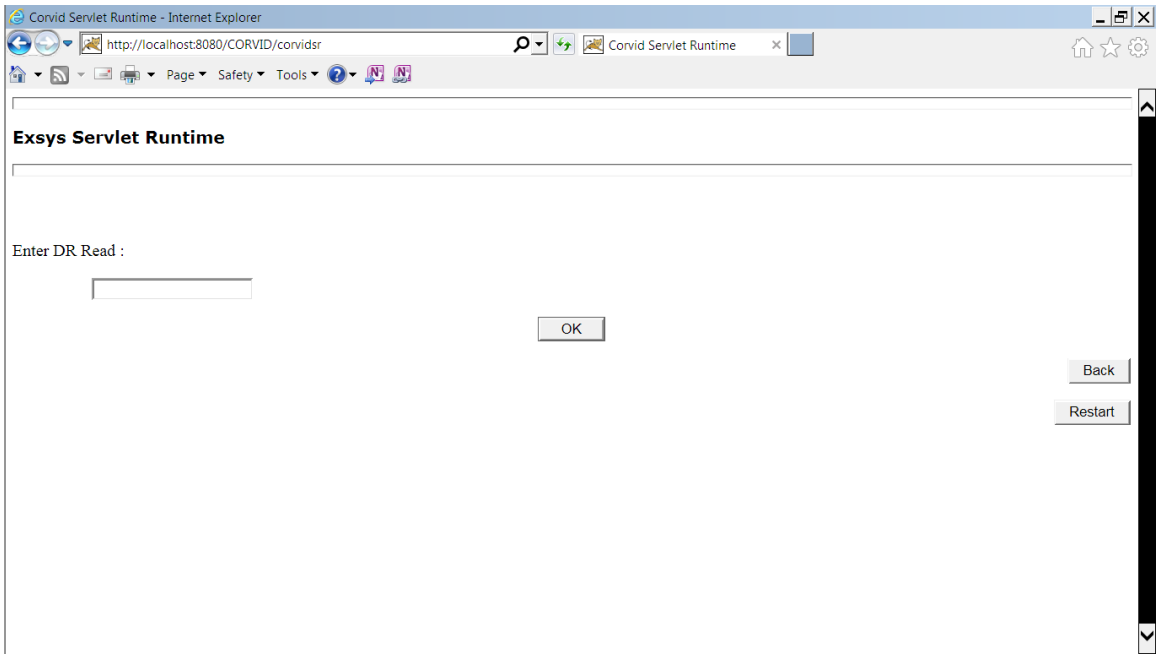
شكل رقم(18.3) يوضح ادخال قراءة AREA



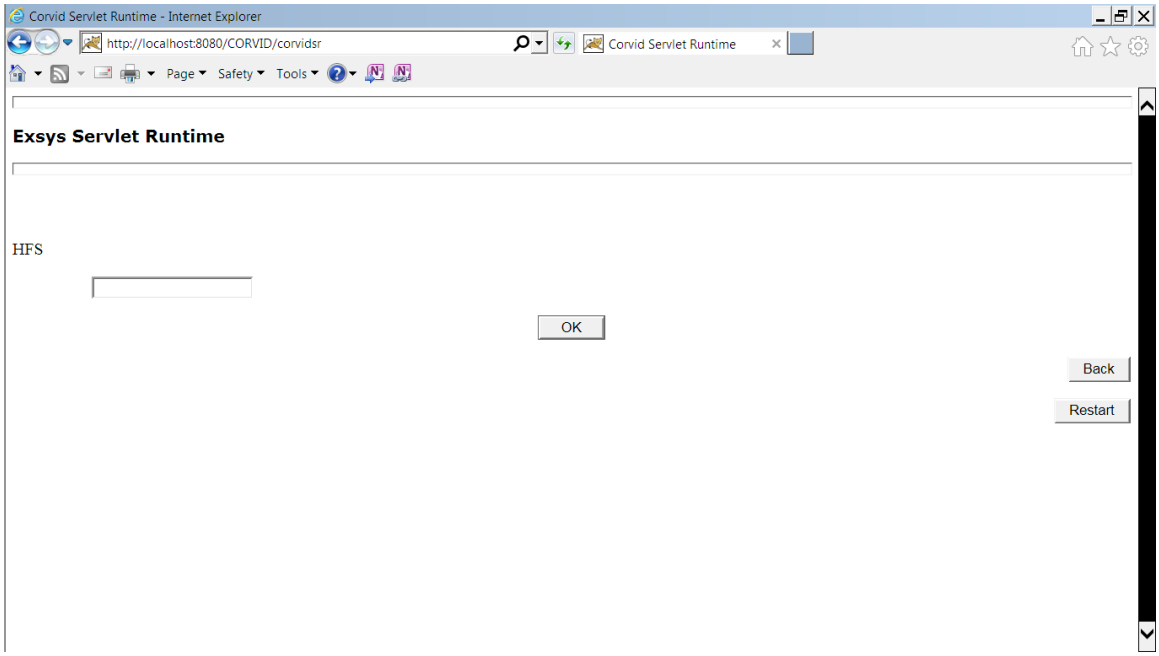
شكل رقم(19.3) يوضح ادخال قراءة DA



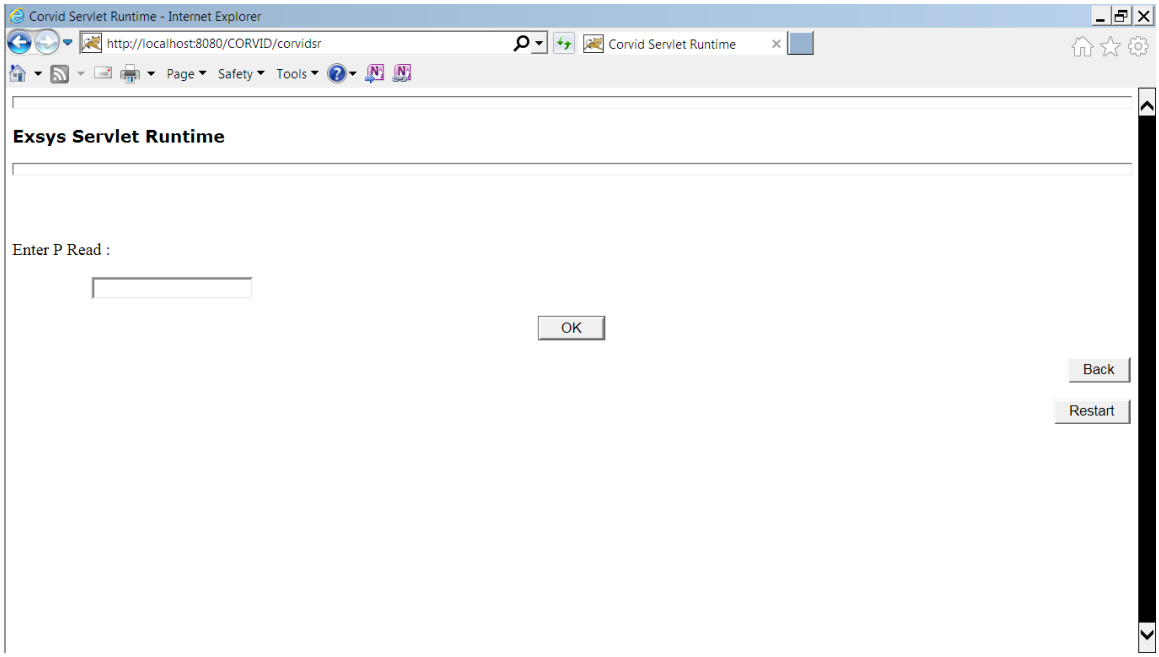
شكل رقم(20.3) يوضح ادخال قراءة A/DA



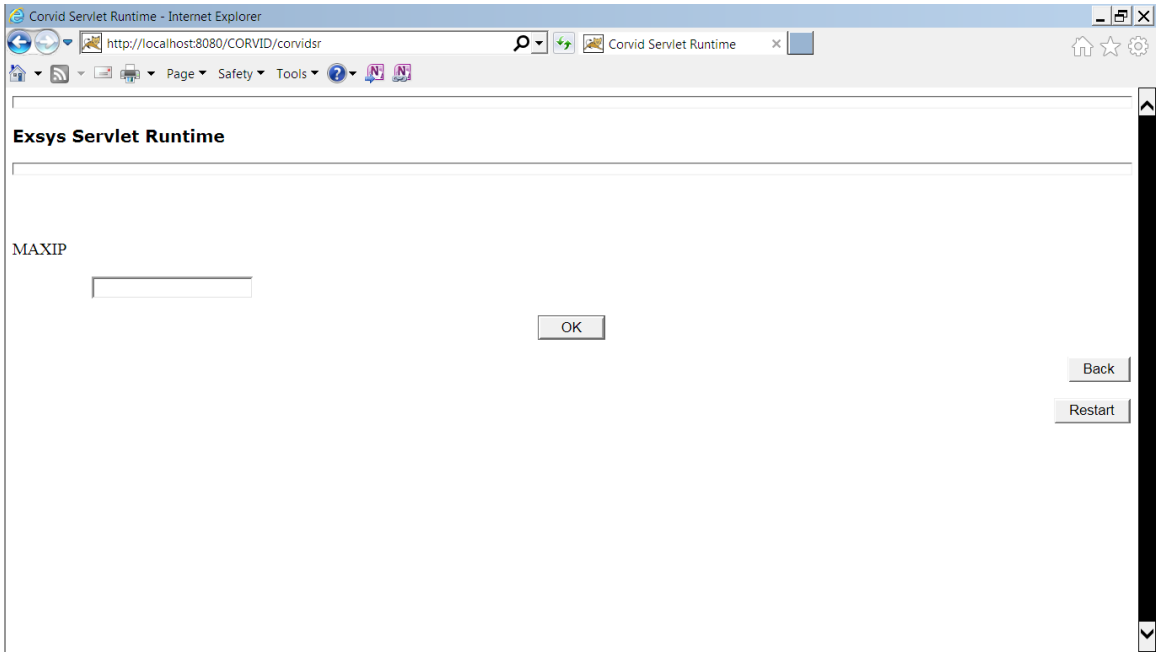
شكل رقم(21.3) يوضح ادخال قراءة DR



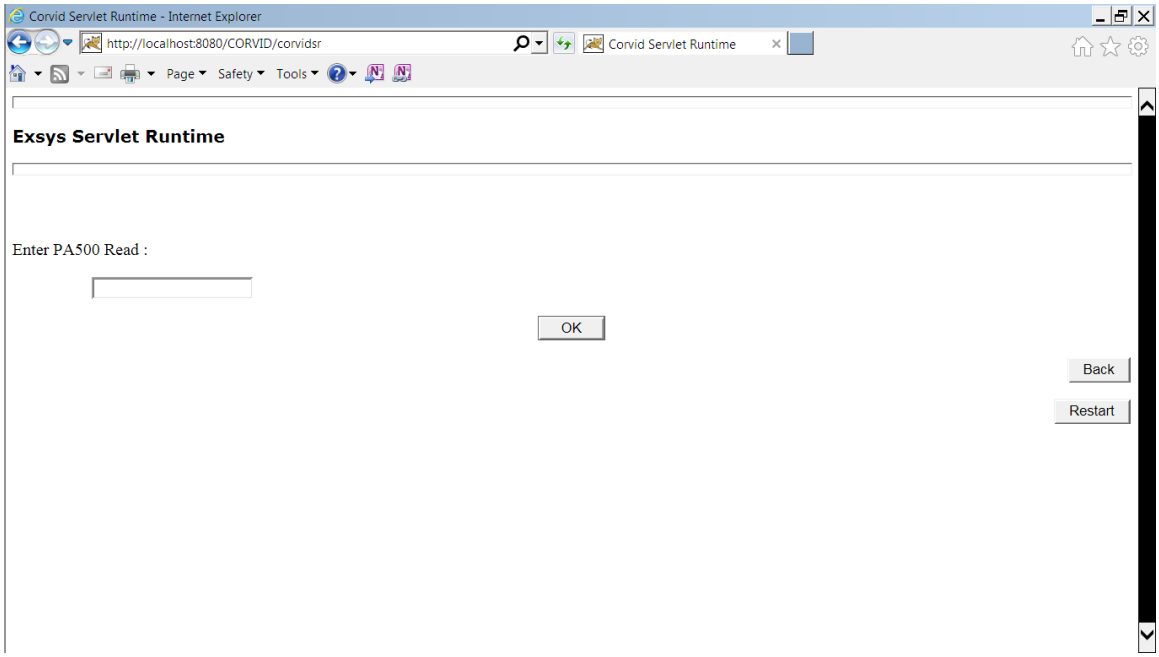
شكل رقم(22.3) يوضح ادخال قراءة HFS



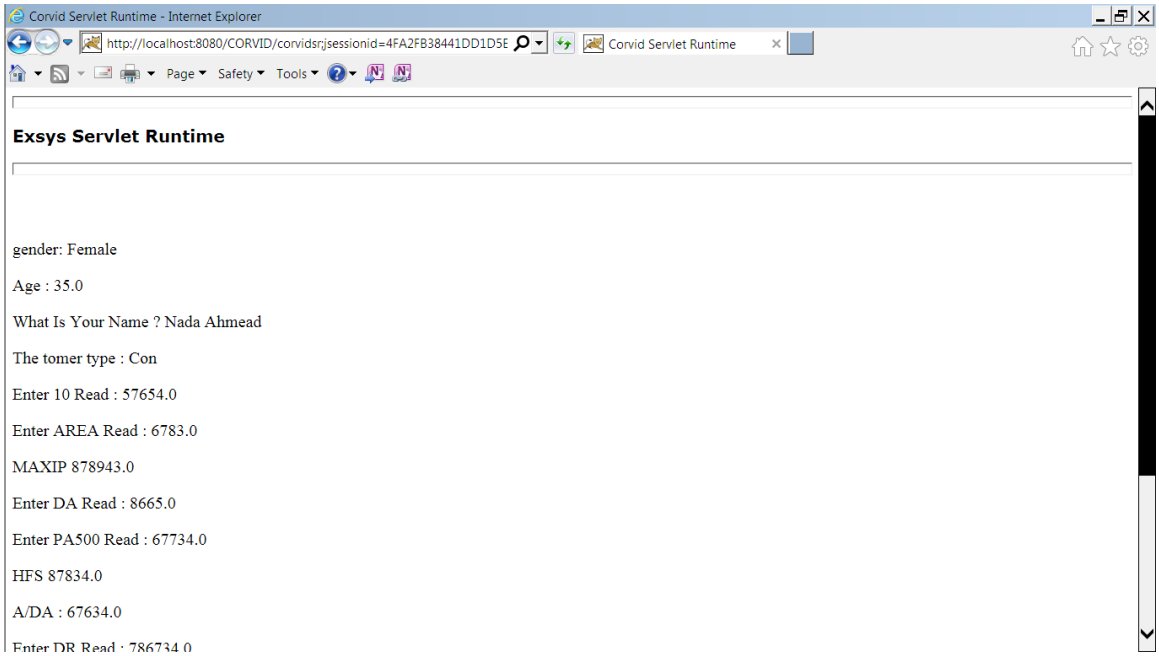
شكل رقم(23.3) يوضح ادخال قراءة P



شكل رقم(24.3) يوضح ادخال قراءة MAXIP



شكل رقم(25.3) يوضح ادخال قراءة PA500



شكل رقم(26.3) يوضح مثال لتنفيذ البرنامج

البـــــاب الرابع

النتائج و التوصيات

1.4 مقدمة:

يحتوي هذا الباب على النتائج التي تم التوصل إليها عند بناء النظام والتوصيات التي يمكن أن تساهم في الاستفادة من هذا النظام وتحقيق الفوائد المرجوة منه , وكذلك الخاتمة و الملاحق.

2.4النتائج:

لقد تم استخدام عدد من التقنيات و اللغات في تطوير و تصميم برنامج طبي يقدم خدماته للأطباء و المرضى عن طريق تشخيص أورام الثدي و تشمل تقنيةData mining المتمثلة في خوازمية C5و تقنية Expert Systemالمتمثلة في ExsysCorvidوتم الدمج بين التقنيتين في نظام هجيني من يوفر خصائص مبتكرة .

3.4التوصيات :

يوصى بتنفيذ النظام في المستشفيات والمراكز الصحية وتطويره وكما يوصي الباحثون بتنفيذ النظام بصورة واسعة بما يسهم بدمج التقنيق الحديثة في الطب وتسهيل العملق التشخيصية.

وأخيرا يمكن القول بان نجاح صناعة التشخيص الالكتروني عبر الحاسوب يجب عمل ما يلي :

1. التعبئة الاجتماعية لدى أفراد المجتمع للتفاعل مع هذا النوع من الأنظمة.
2. تطوير النظام حتى يتيح إضافة خصائص تشخيصية جديدة .
3. استخدام تقنيات بديلة غير المستخدمة .
4. ربط النظام بشبكات الويب وتوفير Communityللتفاعل المباشر بين الأطباء و المرضى .
5. التعديل في واجهة المستخدم الرئيسية حتى يتمكن المريض من تقديم سؤاله.
6. وضع برامج لتدريب الأطباء للاستفادة القصوى من التقنية.
7. تأجير أو استخدام خطوط اتصال منفصلة لضمان الاستخدام الأمثل لهذا النوع من الطب بحيث توفر سرعة , و ذلك في حالة أن تم ربطه بالويب.

4.4الخاتمة:

بحمد الله وعونه تم انجاز هذا العمل المتواضع و الذي كان في مجال من مجالات الحياة التي تخدم الإنسانية وتقدم العون للبشرية وهو مجال تشخيص أورام الثدي الذي يكون بدوره الدافع للمساهمة في الحد من خطورة المرض بالتشخيص المبكر والدقيق لمكافحة سرطان الثدي والاستفادة من التقنيات الحديثة لجعل العملية التشخيصية متوفرة لكل شرائح المجتمع.

نرجو من الله تعالى أن نكون قد وفقنا في هذا العمل ولو بالقليل ليسهم في دفع هذه المسيرة الجديدة الرائدة في المجال الطبي وان تكون التقنيات المستخدمة قد أسهمت في حل وترقية هذا التطبيق والحمد لله من قبل وبعد.

5.4 الملاحق:

Dataset with electrical impedance measurements in samples of freshly excised tissue from the breast[3]

106 instances
10 attributes: 9 features+1class attribute

Six classes of freshly excised tissue were studied using electrical impedance measurements:

		# of cases
Car	Carcinoma	21
Fad	Fibro-adenoma	15
Mas	Mastopathy	18
Gla	Glandular	16
Con	Connective	14
Adi	Adipose	22
		106

Impedance measurements were made at the frequencies: 15.625, 31.25, 62.5, 125, 250, 500, 1000 KHz
These measurements plotted in the (real, -imaginary) plane constitute the impedance spectrum from where the features below are computed.[3]

9 features:

I0	Impedivity (ohm) at zero frequency
PA500	phase angle at 500 KHz
HFS	high-frequency slope of phase angle
DA	impedance distance between spectral ends
AREA	area under spectrum
A/DA	area normalized by DA
MAX IP	maximum of the spectrum
DR	distance between I0 and real part of the maximum frequency point
P	length of the spectral curve[3]

Data Set(1.5)جدول[3]

class	I0	PA500	HFS	DA	AREA	A/DA	MAX IP	DR	P
car	524.79407	0.1874484	0.0321141	228.80023	6843.5985	29.910803	60.20488	220.73721	556.82833
car	330	0.2268928	0.26529	121.1542	3163.2395	26.109202	69.717361	99.084964	400.22578
car	551.87929	0.2324779	0.06353	264.80494	11888.392	44.894903	77.793297	253.7853	656.76945
car	380	0.2408554	0.286234	137.64011	5402.1712	39.248524	88.758446	105.19857	493.70181
car	362.83127	0.2007129	0.2443461	124.91256	3290.4624	26.342127	69.389389	103.86655	424.7965
car	389.87298	0.1500983	0.0977384	118.62581	2475.5571	20.86862	49.757149	107.68616	429.38579
car	290.45514	0.1441642	0.053058	74.635067	1189.5452	15.938154	35.703331	65.541324	330.26729
car	275.67739	0.153938	0.1877974	91.527893	1756.2348	19.187974	39.305183	82.658682	331.5883
car	470	0.2131047	0.2254965	184.59006	8185.3608	44.343455	84.482483	164.12251	603.31572
car	423	0.2195624	0.2617994	172.37124	6108.1063	35.435762	79.056351	153.1729	558.27452
car	410	0.3178245	0.2974041	255.81518	10622.547	41.524303	67.523209	246.74283	508.54036
car	500	0.2272419	0.0509636	219.2955	9819.4496	44.77725	76.8685	207.26664	602.52784
car	438.78016	0.2124066	0.0607375	120.9016	4879.4956	40.359232	80.791779	89.943786	525.42015
car	366.94238	0.2801253	0.2520255	172.74555	7064.8159	40.897237	75.604324	155.32228	471.5882
car	485.66881	0.2302089	0.1340413	253.8937	8135.9684	32.044783	64.855446	245.47053	541.36398
car	390	0.3583161	0.2038545	245.6861	10055.837	40.929612	70.324783	236.49017	477.54836
car	269.49595	0.2075196	0.0383972	80.411085	1963.6052	24.419584	44.740154	66.838309	329.09065
car	300	0.1900664	0.1668535	97.10813	3039.5613	31.300791	51.353973	82.418192	387.07823
car	325	0.2246239	0.2869321	229.21586	5705.3321	24.890651	35.602715	227.26479	462.70301
car	294.47485	0.206647	0.4677482	194.87104	5541.2561	28.435504	36.765797	191.80489	445.5133
car	500	0.1926843	0.1947787	144.68858	3055.013	21.114403	96.56337	107.7511	542.89709
fad	211	0.0539307	0.0942478	30.753443	151.98458	4.9420345	14.268374	27.243124	217.1307
fad	196.85671	0.0200713	0.0907571	28.593126	82.058889	2.8698817	7.9687834	27.661516	200.74934
fad	245	0.1890192	0.0816814	62.902955	1235.9834	19.649051	42.152016	46.690355	292.37624
fad	352.65645	0.1219985	0.0907571	68.527846	1066.1578	15.558024	43.691925	52.792817	382.73319
fad	243.29398	0.039968	0.0670206	68.544778	383.92845	5.6011335	9.9913483	67.816656	263.64076
fad	259.88515	0.0706858	0.0069813	58.243807	465.08727	7.9851797	17.506838	56.340241	267.51745
fad	250	0.0680678	-0.015359	57.172431	652.90135	11.419863	17.776981	55.79127	278.30862
fad	200	0.0376991	0.1172861	42.316675	220.81091	5.218059	10.675764	40.947882	218.03431
fad	355	0.0640536	0.0844739	89.558463	1033.8542	11.543903	27.562662	86.576561	372.03996
fad	272	0.0914553	0.0048869	63.78938	718.94631	11.270627	20.085556	60.690729	286.92022
fad	341.62001	0.0919789	0.074002	85.043029	1370.8381	16.119347	29.028969	79.935197	385.13346
fad	160.32245	0.1769764	0.1633628	37.217124	341.88146	9.1861333	30.889164	20.75991	187.56652
fad	301.3044	0.1097812	0.0356047	64.6162	942.77319	14.590353	29.045677	57.72003	335.76846
fad	155	0.1727876	0.1186824	38.940168	415.1116	10.660242	25.836502	29.134376	184.81704
fad	144	0.1206023	0.0460767	19.64767	70.426239	3.5844576	18.131014	7.5694931	160.37377
mas	178	0.1708677	0.2129302	41.542167	489.44052	11.781776	35.747795	21.162393	215.90593
mas	195	0.1391027	0.2052507	37.462196	328.38405	8.7657449	35.023827	13.29465	232.58763
mas	435.09317	0.076969	0.1612684	123.59754	1342.2776	10.860068	37.384724	117.80804	433.20232
mas	250	0.0471239	0.0139626	70.90704	224.14579	3.1611218	9.1021757	70.3204	232.27763
mas	339.50855	0.0453786	0.0300197	88.629809	331.08039	3.7355422	19.825581	87.62118	307.79319
mas	236	0.1246165	0.2017601	48.45113	236.88011	4.8890524	36.006622	32.41967	244.96515

mas	481.46715	0.0785398	0.016057	79.059162	1154.3382	14.600942	33.929131	71.408439	501.89336
mas	252	0.106116	0.0314159	38.54421	493.79042	12.811014	25.541145	28.86704	280.6583
mas	172.5158	0.1272345	0.0383972	37.543673	192.21815	5.1198546	19.322081	32.189821	174.93377
mas	121	0.1741839	0.0907571	24.43718	144.46651	5.9117505	22.02108	10.594708	141.76618
mas	196.36486	0.1829105	0.1424189	54.58308	843.26253	15.449156	34.149414	42.580865	239.94407
mas	370.39572	0.1047198	0	115.92325	1308.1204	11.284366	31.367031	112.7151	365.97765
mas	260.27752	0.0790634	0.0328122	58.817611	277.25903	4.7138778	17.86867	56.037684	248.62341
mas	544.65435	0.0637045	0.0020944	100.78808	1189.2902	11.79991	29.412222	96.58015	553.35821
mas	310	0.1747075	0.1654572	98.509961	2741.032	27.824923	49.327862	85.27001	388.97781
mas	274.9934	0.1471313	0.1375319	66.457943	1217.4157	18.318588	40.849678	52.421008	327.55864
mas	281.32326	0.2323033	0.4377286	157.88418	5305.1232	33.60136	46.384335	150.91689	398.89555
mas	327	0.1410226	0.0844739	76.212868	1664.6741	21.842428	43.221691	62.771703	379.26184
gla	470.51642	0.127409	0.0663225	150.22402	2657.9104	17.692978	47.560631	142.49647	491.47249
gla	223	0.1240929	0.079587	33.096335	197.01254	5.9526996	30.454874	12.95639	252.48335
gla	152	0.1658063	0.2275909	34.21955	94.354328	2.7573223	31.279278	13.877478	180.60955
gla	303	0.0633555	0.0397935	22.567893	102.50031	4.5418644	21.830841	5.7205006	321.64684
gla	250	0.0870919	0.0928515	29.63751	180.76124	6.0990698	26.142097	13.963263	280.12291
gla	197	0.132645	0.074002	33.460653	409.64714	12.242652	26.992807	19.773813	231.78379
gla	197	0.132645	0.074002	33.460653	409.64714	12.242652	26.992807	19.773813	231.78379
gla	216.41326	0.1178097	0.0684169	53.5996	280.44549	5.232231	22.790535	48.512974	215.37292
gla	178	0.1495747	0.0991347	40.290621	474.4047	11.774569	25.920828	30.845499	209.18074
gla	185	0.1497492	0.0851721	39.891409	361.74798	9.0683179	26.860762	29.49278	210.17872
gla	391	0.0581195	0.0111701	35.780061	265.14979	7.410546	22.131472	28.114244	400.99482
gla	502	0.0652753	0.0279253	53.239433	834.27273	15.670203	33.331142	41.514722	544.03941
gla	176	0.0898845	0.0767945	20.588524	79.705425	3.871352	18.226492	9.5750879	191.99288
gla	145	0.1176352	0.1103048	21.218942	82.455562	3.8859412	20.303082	6.1667152	162.51093
gla	124.12873	0.1319469	0.1089085	20.592633	78.342794	3.8044088	18.462399	9.1212056	134.8927
gla	103	0.1581268	0.2918191	23.754811	78.258474	3.2944262	22.323603	8.120826	124.97856
con	1724.0899	0.0527089	-0.020944	404.12621	3053.9669	7.5569632	71.427589	399.19424	1489.3867
con	1385.6647	0.0923279	0.0893609	202.48004	8785.0287	43.387134	143.09219	143.25778	1524.6092
con	1084.247	0.0734784	0	191.89795	2937.9715	15.310072	66.563629	179.98363	1064.1038
con	649.3694	0.1075123	0.0188496	207.11169	3344.4327	16.147967	50.54747	200.84871	623.90881
con	1500	0.0565487	0.0502655	375.0966	4759.4545	12.68861	78.445976	366.80198	1336.1585
con	770	0.0416283	0.0024346	175.01968	346.09131	1.9774422	25.22233	173.19274	654.79842
con	650	0.0410152	0.1452114	216.81133	427.53407	1.9719176	33.765163	214.16598	528.69923
con	691.97203	0.0260054	0.0865683	190.67669	304.27072	1.5957415	23.975718	189.16333	594.31563
con	1461.7503	0.0408407	0.0495674	391.84602	5574.0028	14.224982	57.23148	387.64399	1428.84
con	1496.7361	0.1038471	0.0823795	640.27595	11071.998	17.29254	108.28692	631.05248	1178.2743
con	1111.8141	0.0987856	0.0712094	386.98709	7659.7421	19.793275	86.025124	377.3045	990.97891
con	1270.6667	0.0788889	0.0656244	555.35232	3612.9683	6.505723	68.78167	551.07647	895.18744
con	1647.9398	0.0809833	0.0865683	576.77038	11852.485	20.549747	111.43591	565.90291	1402.8777
con	1535.8507	0.0884882	0.0041888	637.34967	10814.05	16.967217	96.610397	632.16512	1197.7622
adi	2100	0.0619592	-0.045379	390.48251	16640.724	42.615798	125.90041	380.6447	2073.0287
adi	1800	0.0342085	0.042586	301.06035	4406.1543	14.635452	67.625328	293.36692	1742.3757
adi	2100	0.1216494	0.3776893	450.55167	35671.606	79.173176	436.09964	113.19857	2461.4505

adi	1666.1488	0.0123918	0.0584515	72.93105	1402.2317	19.226813	51.854767	58.595763	1746.5775
adi	1700	0.0439823	0.1051268	120.65301	12331.103	102.20303	120.29732	-9.257696	2212.1782
adi	1949.1182	0.0517983	0.0165007	170.33149	3212.0771	18.857799	101.45585	136.81933	1941.3674
adi	1850	0.0791492	0.0694702	253.62146	13113.203	51.70384	160.06546	196.7305	1916.9854
adi	2350	0.0815069	0.2715732	515.28948	27758.64	53.869992	289.56906	426.23116	2457.6767
adi	1800	0.0919789	0.2052507	362.86332	15021.554	41.397278	217.83397	290.20364	1893.6637
adi	1900	0.0544543	0.1137955	272.61803	7481.5935	27.443502	138.35927	234.8985	1924.5179
adi	1800	0.069115	0.1570796	385.5647	13831.725	35.87394	157.57001	351.89748	1823.0324
adi	1850	0.0733038	0.2254965	325.19154	8644.9842	26.584284	208.73997	249.35349	1908.1772
adi	1650	0.0476475	0.0432842	274.42618	5824.8952	21.225727	81.239571	262.12566	1603.0703
adi	2800	0.0830777	0.1843068	583.25926	31388.653	53.815953	298.58298	501.03849	2896.5825
adi	2329.8401	0.066148	0.3532546	377.25337	25369.04	67.246689	336.07516	171.38723	2686.4353
adi	2400	0.0841249	0.2206096	596.04196	37939.256	63.651988	261.34818	535.68941	2447.7724
adi	2000	0.0671952	0.1242674	330.27165	15381.098	46.571051	169.19798	283.63956	2063.0732
adi	2000	0.1069887	0.1054179	520.22265	40087.921	77.059161	204.09035	478.51722	2088.6489
adi	2600	0.2005383	0.2080432	1063.4414	174480.48	164.07154	418.68729	977.55237	2664.5836
adi	1600	0.0719076	-0.066323	436.9436	12655.342	28.963331	103.7327	432.12975	1475.3715
adi	2300	0.0450295	0.1368338	185.44604	5086.2925	27.427344	178.69174	49.59329	2480.5922
adi	2600	0.0699877	0.0488692	745.47437	39845.774	53.450226	154.1226	729.3684	2545.4197

6.4المراجع:

[1] تنقيب البيانات الأستاذ الدكتور علاء حسين الحمامي - 2008

[2] الذكاء الاصطناعي دليل النظم الذكية ميشيل نيجينفيتسكي - 2004

[3] J. Jossinet (1996) Variability of impedivity in normal and pathological breast tissue. Med. & Biol. Eng. &Comput, 34: 346-350.
JE Silva, JP Marques de Sá, J Jossinet (2000) Classification of Breast Tissue by Electrical Impedance Spectroscopy. Med & Bio Eng& Computing, 38:26-30.

[4] Top 10 algorithms in data mining
Xindong Wu · Vipin Kumar · J. Ross Quinlan · JoydeepGhosh· Qiang Yang ·
Hiroshi Motoda· Geoffrey J. McLachlan · Angus Ng · Bing Liu · Philip S. Yu Zhi-Hua Zhou · Michael Steinbach · David J. Hand · Dan Steinberg
Received: 9 July 2007 / Revised: 28 September 2007 / Accepted: 8 October 2007
Published online: 4 December 2007

[5] http://spss-clementine.software.informer.com/ 26/8/2014 9:00

[6]مرشد طبي خبيرمتعدد الوسائط – بحث تكميلي لنيل درجة بكلاريوس الشرف في الحاسوب ونظم المعلومات-2005

[5] استخدام تقنية التنقيب في البيانات لمرضى القلب- بحث تكميلي لنيل درجة بكلاريوس الشرف في علوم الحاسوب وتقانة المعلومات-2008