جامعة السودان للعلوم والتكنلوجيا

كلية علوم الحاسوب وتقانة المعلومات

التسويق الذكي بإستخدام التنقيب في البيانات

Intelligent Marketing using data mining

بحث مقدم كأحد المتطلبات الحصول على بكالريوس الشرف في علوم الحاسوب.

أكتوبر 2015

بسم الله الرحمن الرحيم

جامعة السودان للعلوم والتكنلوجيا

كلية علوم الحاسوب وتقانة المعلومات

التسويق الذكي بإستخدام التنقيب في البيانات

Intelligent Marketing using data mining

إعداد الطالبات:

1. **سليمى عبد العزيز أحميدي.**
2. **رهام خالد محمد الخير.**

بحث مقدم كأحد المتطلبات الحصول على بكالريوس الشرف في علوم الحاسوب.

توقيع الاستاذ المشرف:

د.هويدا علي عبد القادر

أكتوبر *2015*

الآيـة

قال تعالى: (قل لو كان البحر مداداً لكلمات ربي لنفد البحر قبل أن تنفد كلمات ربي ولو جئنا بمثله مدداً)

صدق الله العظيم

سورة الكهف"109"

الحمدلله

الحمد لله الذي بنعمته تتم الصالحات , وتكتمل الأعمال , الحمد لله عدد ما كان وعدد ما يكون وعدد الحركات والسكون ,الحمد لله في الأولى والأخرة.

ولا نرى في الحمد الا زيادة في العطاء فنحمدك الله على برك الجزيل ونشكر فضلك على توفيقك لنا وأعانتنا في إتمام هذا البحث وتقديمه بالصورة المطلوبة.

ربي أدم صلاتك الكاملة وبركاتك الشاملة وسلامك الأتم على سيد الخلق اجمعين محمد خاتم النبيين صل الله عليه وسلم.

الأهداء

وإلى من تتسابق الكلمات لتخرج معبرة عن مكنون ذاتها
من علمتني وعانت الصعاب لأصل إلى ما أنا فيه
وعندما تكسوني الهموم أسبح في بحر حنانها ليخفف من آلامي

,,, أمـــــي ,,,

إلى سبب وجودي في الحياة ..
لك كل التجلى والاحترام

,,,والدي الحبيب,,,

إلي كل من أضاء بعلمه عقل غيره
أو هدى بالجواب الصحيح حيرة سائليه
فأظهر بسماحته تواضع العلماء
وبرحابته سماحة العارفين .

,,,أساتذتـــــي,,,

إلى من كانوا يضيئون لي الطريق
ويساندوني ويتنازلون عن حقوقهم
لإرضائي والعيش في هناء

,,,إخوتـــــي,,,

الى رفقـــــــاء المسيرة والدرب من فرحنا سويا وبكينا سويا ورسمنـا ملامح أربع سنوات على دفاتر الزمن ولونها لنا القدر بفراق أجساد وبقاء أرواح وذكريات

,,,أصدقائـــــي الأوفياء

شكر وعرفان

لابد لنا ونحن نخطو خطواتنا الأخيرة في الحياة الجامعية من وقفة نعود إلى أعوام قضيناها في رحاب الجامعة مع أساتذتنا الكرام الذين قدموا لنا الكثير باذلين بذلك جهودا كبيرة في بناء جيل الغد لتبعث الأمة من جديد...
وقبل أن نمضي تقدم أسمى آيات الشكر والامتنان والتقدير والمحبة إلى الذين حملوا أقدس رسالة في الحياة إلى الذين مهدوا لنا طريق العلم والمعرفة.
إلى جميع أساتذتنا الأفاضل.

وأخص بالتقدير والشكر:

الدكتورة :هويدا علي عبدالقادر
التي نقول لها بشراك قول رسول الله صلى الله عليه وسلم:
"إن الحوت في البحر ، والطير في السماء ، ليصلون على معلم الناس الخير".

وكذلك نشكر كل من ساعد على إتمام هذا البحث وقدم لنا العون ومد لنا يد المساعدة وزودنا بالمعلومات اللازمة لإتمام هذا البحث ونخص بالذكر:

الأستاذة: أبتهال مصطفى.

نسأل الله أن يوافيهم بأحسن ما عملوا ويجزيهم خير الجزاء.

المستخلص

يعتبر التسويق من أهم أسباب نجاح المؤسسات , لذلك لابد أن يتم التسويق بطريقة فعالة بإستهداف فئة معينة على حسب إحتياجاتهم من الموارد المسوقة من قبل المؤسسة , وذلك يحفظ زمن المستخدم وموارده بإطلاعه على الإعلانات التي تهمه فقط , بهذه الطريقة يحفظ موارد المؤسسة ايضا .

في هذا البحث أستخدم التنقيب عن البيانات لجعل التسويق موجه وأكثر كفاءة وذلك بتصنيف بيانات عملاء البنك المخزنة مسبقا على حسب خاصية محددة هي إستجابة العميل في الحملات التسويقية السابقة . ركز البحث على إستخدام واحدة من خوارزميات التصنيف وهي شجرة القرار . مخرجات الخوارزمية عبارة عن قواعد تم إستخدامها في نموذج التصنيف للبيانات الجديدة وكانت النتائج بدقة (92.06%) , هذا النموذج يصلح للبنوك لتصنيف عملائهم بنفس الخصائص الواردة في البيانات .

Abstract

Marketing is one of the most important reasons for the success of the enterprise, so it must be done with an effective way of targeting a particular category of customers according to their needs from Marketer resources by the enterprise.This saves the user time and resources to informing him about relevant ads only.

This save enterprise resource as well.

In this research data mining is used to make marketing more efficient and directed through classifying the bank's customers data which is already stored according to specific property , that is the client's response to previous marketing campaigns . the research focused on using one classification algorithm ,The decision tree algorithm.

The outputs are rules that were used in the classification model for new data. The accuracy of the results was 92.06 The model is suitable for use by banks to classify their customers using the same properties found in the data .

المصطلحات الموحدة

شرح المصطلح	المصطلح
التنقيب في البيانات	Data mining
اكتشاف المعرفة في قواعد البيانات	Knowledge discovery in database
المعالجة المسبقة للبيانات	data pre_processing
تصنيف البيانات على حسب خاصية محددة	Classification
الطريقة المتبعة لتقسيم البيانات	decision tree
من خوارزميات شجرة القرار	c&r tree
من برامج تنقيب البيانات(tool)	SPSS Clementine

فهرست الأشكال

رقم الشكل	موضوع الشكل	رقم الصفحة
1-2	خطوات إكتشاف المعرفة من قواعد البيانات	4
1-3	تحميل البيانات على البرنامج Clementine	22
2-3	تحميل البيانات في جدول	**22**
3-3	تصنيف البيانات على حسب خاصية محددة(y)	**23**
4-3	إختيار الخوارزمية المستخدمة	**23**
5-3	نتيجة الخوارزمية المستخدمة	**24**
6-3	خوارزمية اخرى للمقارنة	**24**
7-3	نتائج تحليل البيانات الناتجة من الخوارزمية	**25**
8-3	دقة الخوارزمية المستخدمة	**25**
9-3	نموذج البيانات المخرجة	**26**
10-3	جدول بيانات العملاء في قاعدة البيانات	**27**
11-3	الواجهة الرئيسية للنموذج	**28**
12-3	واجهة تسجيل الدخول	**29**
13-3	واجهة الإختيار.	**30**
14-3	واجهة إضافة العملاء.	**31**
15-3	واجهة معرفة إستجابة العميل.	**32**
16-3	واجهة عرض بيانات العملاء.	**33**

فهرست المحتويات

الموضوع	رقم الصفحة
الآية	أ
الحمد	ب
الإهداء	ت
شكر وعرفان	ث
المستخلص	ج
Abstract	ح
المصطلحات الموحدة	خ
فهرست الأشكال	د
فهرست المحتويات	ذ
الباب الأول	
1.1. مقدمة	1
1.2. مشكلة البحث	1
1.3. أهمية البحث	1
1.4. أهداف البحث	1
1.5. حدود البحث	2
1.6. منهج البحث	2
1.7. مصدر البيانات	2
1.8. محتويات البحث	2
الباب الثاني	
2.1.1.1. تعريف	3
2.1.1.2. التنقيب عن البيانات واكتشاف المعرفة	3
2.1.1.3. مهام التنقيب عن البيانات	4
2.1.1.4. المهام الجوهرية للتنقيب	5
2.1.1.5. التصنيف والتنبؤ	6
2.1.1.6. إعداد البيانات للتصنيف والتنبؤ	6
2.1.1.7. مقارنة طرق التصنيف والتنبؤ	7
2.1.1.8. تقنيات التصنيف	7
2.1.1.9. التصنيف بموجب شجرة القرار التعريفية	8
2.1.2.1.الإدارة الذكية	10
2.1.2.2. التسويق	11
2.1.2.3. مفاهيم أساسية	12

2.1.2.4. التسويق الذكي	12
2.1.2.5. استراتجيات التخطيط للتسويق الذكي	13
2.1.2.6 . التسويق المصرفي	14
2.1.2.7. التسويق بإستخدام التنقيب في البيانات	15
2.1.2.8تحليل وإدارة علاقات العملاء	15
2.2. الدراسات السابقة	16
البـــــــاب الثالث	
3.1. مقدمة عن التطبيق	19
3.2. التقنيات المستخدمة	19
3.2.1. برنامج(SPSS Clementine)	19
3.2.2. HTML	19
3.2.3. PHP	20
3.2.4. MySQL	20
3.3. وصف البيانات	20
3.4. التطبيق على البيانات	21
3.5. بناء واجهات النموذج	27
3.5.1. تطبيق قاعدة البيانات	27
3.5.2. الواجهات	28
البـــــــــــاب الرابع	
4.1. النتائج	34
4.2. التوصيات	35
الخاتمة	36
المراجع	37

الباب الأول

المقدمة

1.1. مقدمة:

التنقيب في البيانات : يستخدم هذا المصطلح لتحليل كميات كبيرة من البيانات من أجل إكتشاف أنماط وقواعد ذات مغذى .الهدف من جمع البيانات هو السماح للشركات بتحسين التسويق وعمليات دعم قرار العملاء من خلال الفهم الأفضل للعملاء .[1]

سابقا كانت عملية التسويق غير موجهة اي بمعني انه يتم استهلاك موارد كثيرة في الاماكن غير الصحيحة . قدمت لنا التقنيات المتطورة القدرة علي جمع كميات كبيرة من البيانات في مختلف المجالات ، تقدم هذه البيانات القوة والجهد لاكتشاف معلومات و معرفة مهمة لم نستطيع معرفتها سابقا .التنقيب عن البيانات يساعدنا علي استخلاص معارف مهمة من أحجام كبيرة من البيانات باستخدام ادوات حديثة.

2.1. مشكلة البحث:

في كثير من الأحيان عندما يتم التسويق لشي معين يتم بطريقة غير فعالة او غير موجهة بمعنى ان تقوم المؤسسة بإرسال اعلانتاتها لكل العملاء بصرف النظر عن أهمية الأعلان للعميل . وذلك قد يجعل الجهه التي تقوم بالتسويق تهدر مواردها ووقتها وجهدها في اماكن غير مناسبة وكذلك يؤدي الي ضياع زمن المستخدم بإطلاعه علي اعلانات غير مفيدة له .

3.1. أهمية البحث:

هذا البحث مهم في حفظ موارد الجهة المقدمة للتسويق من جهد ومال ووقت ، ومن ناحية اخرى فهو مهم للمستخدم بحيث انه يجعله يطلع على الاعلانات المفيدة له فقط وذلك بتصنيف العملاء بناء على البيانات المخزنة مسبقا للعملاء .

4.1. أهداف البحث:

1-ترشيد موارد المؤسسات باستخدام نظم التسويق الذكي.

2- حفظ زمن المستخدم وموارده .

3-جعل التسويق فعال بإتباعه طرق فعالة.

4-الاستفادة من البيانات المخزنة سابقاً في حملات التسويق.

5.1. حدود البحث:

بناء نظام يستهدف تسويق خدمات البنوك كنموذج .

6.1. منهج البحث:

وصفي,تطبيقي.

7.1. مصدر البيانات:

Uci machine leaning repository

8.1. محتويات البحث:

1-الباب الاول (المقدمة).

2-الباب الثاني

الفصل الأول: الإطار النظري ويحتوي على

القسم الأول : التنقيب عن البيانات

القسم الثاني: الإدارة الذكية والتسويق

الفصل الثاني: الدراسات السابقة

3-الباب الثالث(التطبيق).

4-الباب الرابع (النتائج والتوصيات).

الباب الثاني

الإطار النظري والدراسات السابقة

1.1.2 التنقيب عن البيانات:

1.1.1.2 تعريف:

إن التنقيب عن البيانات هو عملية استكشاف لبيانات مفيدة في مصادر بيانات ضخمة. يتم تطبيق تقنيات التنقيب عن البيانات في قواعد بيانات ضخمة بهدف إيجاد نماذج غريبة ومفيدة قد تبقى غير معروفة بدونها . كما انها تقدم إمكانيات لتوقع خرج مشاهدة مستقبلية.

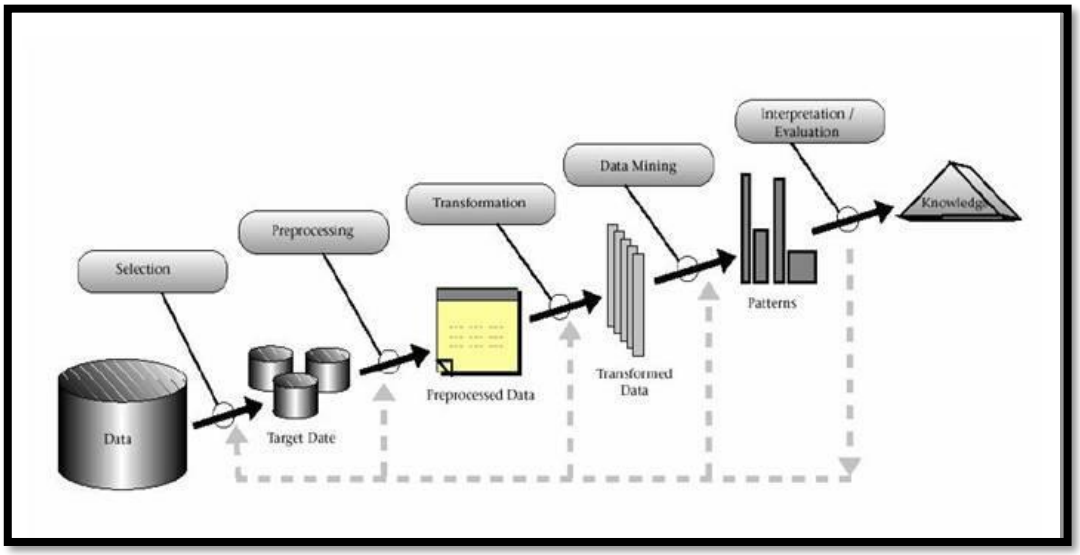
2.1.1.2 التنقيب عن البيانات واكتشاف المعرفة:

يعتبر التنقيب عن البيانات جزءاً مكملاً من اكتشاف المعرفة في قواعد البيانات (Knowledge discovery in database) وهي العملية الكلية التي يتم فيها تحويل البيانات الخام الي معلومات مفيدة . تتكون هذه العملية من سلسلة من خطوات التحويل بدءاً بالمعالجة المسبقة للبيانات (data pre-processing) الي المعالجة الاحقة (post processing)لنتائج التنقيب عن البيانات[6] . وتتكون من تسلسل متكرر من الخطوات التالية:-

1. تنظيف البيانات (Data cleaning):لإزالة الضجيج والبيانات غير متناسقة.
2. تكامل البيانات (Data integration):حيث يمكن الجمع بين مصادر بيانات متعددة.
3. اختيار البيانات(Data selection): حيث يتم إختيار البيانات ذات الصلة لهذه المهمة تجمع من قاعدة البيانات.
4. تحول البيانات(Data transformation) :حيث يتم تحويل البيانات إلى أشكال مناسبة.
5. التنقيب في البيانات(Data mining):عملية أساسية لتطبيق طرق ذكية لاستخراج أنماط البيانات.
6. تقييم الأنماط(Pattern evaluation) :لتحديد التعرف على أنماط مثيرة للاهتمام لتمثيل المعرفة.
7. تمثيل المعرفة(Knowledge presentation) :تقنيات لتمثيل و تصوير المعرفة تستخدم لتقديم المعرفة المفيدة للمستخدم.

الخطوات من 1 إلى4 تمثل أشكال مختلفة لتجهيز البيانات،حيث يتم إعدادها للتنقيب . فتشمل دمج البيانات الواردة من عدة مصادر , وتنقية البيانات لإزالة التشويش (noise) . وباعتبار أن هنالك طرق عديدة لجمع وتخزين البيانات , فإن المعالجة المسبقة للبيانات قد تكون أكثر الخطوات إجهاداً واستهلاكاً للوقت من بين خطوات عملية اكتشاف المعرفة .

استخراج البيانات هو عملية اكتشاف أنماط مثيرة للاهتمام ومعرفة من كميات كبيرة من البيانات .ويمكن أن تشمل مصادر البيانات قواعد البيانات ومستودعات البيانات، الويب،ومصادر اخرى.[2]



شكل(1-2) خطوات إكتشاف المعرفة من قواعد البيانات.

3.1.1.2 مهام التنقيب عن البيانات:

تنقسم مهام التنقيب عن البيانات إلي فئتين أساسيتين:

مهام تنبؤية(predictive Tasks):

الغاية من هذه المهام هي التنبؤ بقيمة سمة معينة استناداً إلي قيم سمات أخرى . تُعرف السمة التي سيتم التنبؤ بها بمتحولات الهدف (target)أوالمتحول التابع(dependent) في حين تُعرف السمات المستخدمة لإجراء التنبؤ بالمتحولات التفسيرية (explanatory) أو المتحولات المستقلة(independent).

مهام وصفية (descriptive tasks):

الغاية هنا هي اشتقاق أنماط (ارتباطات (correlation) نزعات(trends)عناقيد (cluster) مسارات منحنية (trajectories) و التشوهات (anomalies)) تلّخص العلاقات الضمنية في البيانات . غالباً ما تكون المهام الوصفية للتنقيب عن البيانات تفسيرية بطبيعتها وتتطلب عادة تقنيات معالجة لاحقة للتحقيق من صحة النتائج وشرحها.

4.1.1.2 المهام الجوهرية للتنقيب:

1- النمذجة التنبؤية(predictive modeling) :

تشير إلي مهمة بناء نموذج للمتحول الهدف كتابع للمتحولات التفسيرية هناك نوعان من النمذجة التنبؤية : التصنيف (classification) الذي يستخدم من أجل متحولات الهدف المنفصلة (discrete)والإنحدار الذي يستخدم من أجل متحولات الهدف المستمرة . الغاية من كلتا المهمتين هي الحصول على نموذج يجعل الخطأ بين القيم المتنبأ بها والقيم الفعلية للمتحول الهدف صغيراً.

2- تحليل الإقتران (association analysis) :

يستخدم لإكتشاف الأنماط التي تصف بقوة الميزات المقترنة في البيانات . يتم عادة تمثيل الأنماط المكتشفة بشكل قواعد تضمين أو مجموعات جزئية للميزات. وبسبب الحجم الأساسي لفضاء البحث فإن الغاية من تحليل الإقتران هي استخلاص أكثر الأنماط أهمية بطريقة فعالة.

3- تحليل العناقيد(cluster analysis) :

يلتمس إيجاد جماعات من المشاهدات وثيقة الصلة ببعضها بحيث تكون المشاهدات التي تنتمي إلي نفس العنقود مشابهة لبعضها البعض أكثر من شبهها بالمشاهدات التي تنتمي إلي عناقيد أخرى . تم إستخدام العنقدة لتجميع مجموعات من الزبائن الذين يجمعهم رابط وإيجاد مناطق من المحيط لها تأثير هام على مناخ الأرض ولضغط البيانات.

4- اكتشاف التشوهات(anomaly detection) :

وهو مهمة تحديد المشاهدات التي تكون مميزاتها مختلفة بشكل ملحوظ عن باقي البيانات . تُعرف هذه المشاهدات بالتشوّهات (anomalies) أو الشواذ(outliers) . أما الغاية من خوارزمية اكتشاف التشوّهات فهي اكتشاف التشوّهات الحقيقية وتجنب التسمية الخاطئة للكائنات الطبيعية على أنها غير سوية . بمعنى أنه يجب أن يكون للكشف الجيد عن التشوّهات معدّل عالٍ للكشف (detection rate) ونسبة منخفضة للإنذار الخاطئ (false alarm rate) من التطبيقات التي يستخدم فيها اكتشاف التشوّهات هنالك اكتشاف الاحتيال (fraud)والتطفل على الشبكة(network intrusion) وأنماط الأمراض واضطراب النظام البيئي .[6]

5.1.1.2 التصنيف والتنبؤ(Classification and Prediction):

هو تحليل للبيانات من أجل إستخراج نماذج تصف فئات بيانات هامة مثل هذه النماذج تدعى المصنفات (classifiers) , على سبيل المثال، يمكننا أن نبني نموذج التصنيف طلبات الحصول على القروض المصرفية إما آمنة أو خطرة.

هذا التحليل يمكن أن يساعد في توفر فهما أفضل للبيانات ككل.لقد تم اقتراح العديد من أساليب التصنيف من قبل الباحثين في تعلم الألة والتعرف على الأنماط والاحصائيات. من أمثلة المصنفات :شجرة القرار , المصنفات القاعدية (rule-based), والشبكات العصبية(neural network), وآلة متجه الدعم(naïve bayes classifiers) , ومصنفات بيز المبسطة . تستخدم كل تقنية خوارزمية تعلم (learning algorithm) لتعريف نموذج يلائم افضل ما يمكن العلاقة بين مجموعة السمات وتسمية الصنف للبيانات المدخلة . يجب أن يلائم النموذج الذي تولده الخوارزمية بيانات الدخل بشكل جيد ويتنبأ بشكل صحيح تسميات الأصناف لتسجيلات لم يصادفها من قبل .لذلك فأن الهدف الاساسي من خوارزمية التعلم هو بناء نماذج يمكن تعميمها, بمعنى أن هذه النماذج يمكن أن تتنبأ بشكل دقيق بتسميات أصناف تسجيلات غير معروفة مسبقا. التصنيف له تطبيقات عديدة ، بما في ذلك الكشف عن الغش ، استهدافا لتسويق ، والتنبؤ الأداء، والتصنيع، والتشخيص الطبي.

6.1.1.2 إعداد البيانات للتصنيف والتنبؤ(Preparing the Data for Classification and Prediction):

يمكن تطبيق الخطوات التالية لتجهيز البيانات للمساعدة في تحسين دقة وكفاءة وقابلية العملية للتصنيف أوالتنبؤ:

تنظيف البيانات (Data cleaning) : هذا يشير إلى تجهيز البيانات من أجل إزالة أو تقليل الضوضاء ومعاملة القيم المفقودة (على سبيل المثال، عن طريق استبد القيمة مفقودة مع القيمة الأكثر شيوعا لهذه السمة، أو مع القيمة الأكثر احتمالا استنادا إلى الإحصائيات). ورغم أن معظم خوارزميات التصنيف لها بعض الآليات للتعامل مع بيانات صاخبة أو مفقودة، ويمكن لهذه الخطوة أن تساعد في الحد من الارتباك أثناء التعلم.

تحليل صلة(Relevance analysis):كثير من السمات في البيانات قد تكون زائدة عن الحاجة .ويمكن استخدام تحليل الارتباط لتحديد ما إذا كان أي اثنين من سمات معينة ترتبط إحصائيا.

تحويل البيانات والحد منها(Data transformation and reduction): يمكن ان تتحول البيانات عن طريق التطبيع(normalization)، وخاصة عندما تستخدم الشبكات العصبية أو التي تنطوي على أساليب قياس المسافات في خطوة التعلم, التطبيع يشمل رفع كل القيم لسمة معينة بحيث تقع ضمن نطاق معين صغير. ويمكن أيضا أن تتحول البيانات عن طريق تعميم المفاهيم ذات مستوى أعلى. ويمكن استخدام مفهوم التسلسل الهرمي لهذا الغرض. وهذا مفيد بشكل خاص للسمات بقيمة مستمرة.

7.1.1.2 مقارنة طرق التصنيف والتنبؤ(Comparing Classification and Prediction Methods)

ويمكن مقارنة أساليب التصنيف والتنبؤ وتقييمها وفقا للمعايير التالية:

الدقة (Accuracy):

دقة المصنف على مجموعة اختبار معين هو النسبة المئوية للمجموعة اختبار الصفوف التيتصنف بشكل صحيح عن طريق المصنف .تتم مقارنة كل اختبار الصفوف (tuple) مع الطبقة(class)التي تم التنبؤ بها من قبل المصنف لتلك الصفوف(tuple) .

التدرجية(Scalability):

هذا يشير إلى القدرة على بناء المصنف أو تنبؤ معين بكفاءة كميات كبيرة من البيانات.

التفسير(Interpretability):

هذا يشير إلى هذا مستوى الفهم والبصيرة التي يتم توفيرها من قبل المصنف أو التوقع. التفسير غالبا ما يكون ذاتي وبالتالي من الصعب التقييم.[5]

8.1.1.2 تقنيات التصنيف(classification techniques):

توجد تقنيات عديدة في مجال التصنيف وهي كما يلي:

1. الطرق المعتمدة على شجرة القرار(decision tree base method).
2. الطرق المعتمدة على قاعدة(rule base method).
3. تصنيف بايزان(Bayesian classification).
4. طريقة الشيكات العصبية(neural network approach).
5. التفكر المعتمد علي الذاكرة (memory based reasoning).
6. أسناد الماكنات الموجهة(support vector machines).
7. الخوارزميات الجينية(genetic algorithm). [7]

9.1.1.2 التصنيف بموجب شجرة القرار التعريفية(Classification by Decision Tree Induction):

شجرة القرار هي عبارة عن هيكل لمخطط انسيابي يشبه الشجرة ، حيث تحتوي على ثلاثة انواع من العقد :

عقدة داخلية (internal nodes) : لكل منها تفرع قادم واحد وتفرعين خارجين منها أو أكثر و ترمز الى اختبار السمة , ويمثل كل فرع نتائج الاختبار.

عقدة ورقية (leaf node) :لكل منها تفرع قادم واحد فقط وليس لها أي تفرع خارج, وتحمل تسمية الطبقة.

عقدة الجزر (root node): هي العقدة الأكثر علوا في الشجرة وليس لها تفرع قادم .

إن هدف أشجار القرار هو تنبؤ أو توضيح الاستجابات على المتغير المعتمد الحاسم,إن المرونة فيها جعلها خيار تحليلي جذاب جدا لكن ليس معنى هذا أن إستخدامها أكثر من الطرق التقليدية . كتقنية إستبيان عندما تفشل الطرق التقليدية تبقى أشجار القرار في رأي معظم الباحثين لا نظير لها.[6]

- **خصائص أشجار القرار:**

1. الطبيعة الهرمية.

2. المرونة: وذلك لقدرتها على فحص تأثير متغيرات المتنبأ كل واحد في وقت, بدلا من فحصها جميعا في وقت واحد , وقدرتها لإنجاز تجزأة متكررة, فحص تأثيرات المتنبأين واحدة في كل وقت, له تأثير في تغيير أنواع المتنبأين التي يمكن تحليلها.

- **التصنيف بإستخدام أشحار القرار:**

أن العقدة في الشجرة هي عبارة عن فحص لبعض الصفات.أما الفرع فهو قيمة محتملة للصفة. ولعمل التصنيف فيجب عمل الاشياء التالية:

- أبدأ من الجذر.
- أفحص الصفة.
- تحرك إلى اسفل فرع في الشجرة.

تتكون مشكلة التصنيف من مجموعة من القيود, وكل قيد يتكون من مجموعة من الصفات وكذلك مؤشر صنف. يتم بناء نموذج دقيق لكل صنف إعتمادا على مجموعة الصفات. بعد ذلك يتم إستخدام النمذج لتصنيف بيانات مستقبلية والتي تكون مؤشرات أصنافها معروفة.

يتم إستخدام نموذج شجرة القرار وذلك للاسباب التالية:

- الحصول على دقة مشابهة وفي بعض الاحيان أفضل عند مقارنتها مع النماذج الاخرى.
- يكون سريع نسبيا مقارنة مع نماذج التصنيف الأخرى.
- بسيط وسهل الفهم.
- يمكن تحويله إلى ابسط وأسهل لفهم قواعد التصنيف.[7]

- **بناء شجرة القرار:**

هناك العديد من اشجار القرار التي يمكن بناؤها من خلال مجموعة من السمات المعطاة , في حين أن بعض الاشجار اكثر دقة من غيرها فإن إيجاد الشجرة المثالية غير محتمل بسبب الحجم الهائل لفضاء البحث ومع ذلك تم تطوير خوارزمية فعالة لبناء شجرة قرار دقيقة ,تستخدم هذه الخوارزميات خطة إستراتيجية وذلك ببناء شجرة قرار من خلال وضع سلسلة من القرارات المثلى حول السمة التي تستخدم لتصنيف البيانات . تعتبر خوازمية هنت واحدة من هذه الخوارزميات وهي

أساس للعديد من خوارزميات بناء شجرة القرار الموجودة بما فيها (ID3,C4.5,CART,C&R, tree).

- **خوارزمية هنت:**

تنمو شجرة القرار في هذه الخوارزمية بطريقة تكرارية من خلال تقسيم تسجيلات التدريب الى مجموعات جزئية متتابعة أكثر تجريدا .لتكن Dt مجموعة تسجيلات التدريب المقابلة للعقدة t ولتكن y={y1,y2,…,yc} تسميات الاصناف فيما يلي التعريف التكراري لخوارزمية هنت :

الخطوة الأولى : إذا كانت جميع التسجيلات في Dt تنتمي الى نفس الصنف Yt فإن t تكون عقدة ورقية تحمل التسمية Yt .

الخطوة التانية : إذا كانت Dt تسجيلات تنتمي الى أكثر من صنف واحد عندها يتم إختيار شرط إختبار سمة لتقسيم التسجيلات الى مجموعات جزئية أصغر .يتم إنشاء عقدة إبن من أجل كل خرج لشرط الأختبار ويتم توزيع التسجيلات في Dt على الأبناء إستنادا على المخرجات يتم بعدها تطبيق الخوارزمية بشكل تكراري على كل عقدة إبن.[6]

2.1.2 الإدارة الذكية والتسويق:

1.2.1.2 الادارة الذكية:

الإدارة الذكية للأعمال تشير إلى التطبيقات التي تستخدم لجمع البيانات والمعلومات عن عمليات المؤسسة وتحليلها. وتساعد الإدارة الذكية في معرفة العوامل التي تؤثر على أنشطتها التجارية ، وبالتالي تساعد الشركات على اتخاذ قرارات أفضل . الإدارة الذكية للأعمال تمكن المؤسسات من التنافس بفعالية أكبر.

تساعد الإدارة الذكية المؤسسات على تحليل الاتجاهات المتغيرة في حصص السوق، التغير في سلوك المستهلك، أنماط الإنفاق، وتفضيلات المستهلكين.

الإدارة الذكية تساعد المؤسسة في التخطيط الاستراتيجي وذلك اعتمادا على مفهوم أساسي وهو أن كل نشاط بالمؤسسة يجب أن يخدم شخص ما بشكل مباشر أو غير مباشر وهذا يعني أن القائمين على العمل في حاجة إلى معرفة من هم المستفيدون وماذا يحتاجون لتلبية رغباتهم .

2.2.1.2 التسويق:

التسويق هو مجموعة من و الأنشطة التي تقوم بإكتشاف رغبات العملاء وإشباعها وتحقق للمؤسسة الربحية خلال فترة مناسبة .

من الناحية المجتمعية، التسويق هو الرابط بين احتياجات المجتمع وبين الاستجابة لأنماط الاقتصاد من خلال توصيل قيمة منتج أو خدمة إلى العملاء. أما التسويق كعلم، فهي عملية تحديد السوق المستهدف من خلال القيام بتحليل السوق وفهم توجهات العملاء وتقديم قيمة عالية لهم.

❖ **يتألف النشاط التسويقي من أربعة عناصر أساسية حددها "جيروم ماكارتي ":-**

1-المنتج:

يعني مايتم عرضه في السوق ، الخدمات التي يتحصل عليها المشتري .

2-التوزيع او المكان:

الطريقة التي تجعل المنتج في متناول المشتري ووصوله إلى السوق المستهدف.

3-الترويج:

يعني نشاطات الاتصال مثل الإعلان عن المنتج و السوق المستهدف المتواجد به المنتج .

4- السعر:

يعني سعر المنتج بالإضافة إلى الرسوم الأخرى مقابل التوصيل والتصريح .[8]

ثم جاء عملاق التسويق كوتلر وأضاف عنصرين أساسيين آخرين وهما :

1-السياسات:

السياسات المتبعة في الدولة تؤثر على التسويق وبالتالي الإنتاجية.

2-الراي العام:

راي المجتمع يؤثر على إنتاجية المنتج لذلك لابد أن يكون محبب له.

3.2.1.2 مفاهيم أساسية:

هناك بعض المفاهيم الأساسية والتي لا بد من أي مهتم بمجال التسويق أن يعرفها:

التسويق لا يعني الإعلان:

النظرة إلى التسويق والبيع على أنهما شيء واحد هذا غير صحيح ، وذلك لان البيع جزء من التسويق غير أن التسويق يشمل أشياء أكثر من البيع" ويمكننا أن نشبه البيع على أنه القشرة الخارجية للتسويق.

التسويق لا يعني البريد المباشر:

معظم الشركات تحتاج إلى وفرة من أسلحة التسويق الأخرى لدعم البريد المباشر ولإنجاح البريد المباشر حتى تستطيع الشركات الحصول على الحجم الذي تريده من التجارة .

التسويق لا يعني المنشورات الدعائية :

المنشور الدعائي هو جانب مهم من جوانب التسويق ولكن لوحده لا يعني التسويق .

التسويق ليست بصناعة ترفيهية:

التسويق ليس مقصوداً به الترفيه.

التسويق ليس مسرحاً للفكاهة:

الفكاهة اذا تكررت في الحملات التسويقية تعوق المفهوم الأساسي وراء نجاح التسويق .

التسويق لا يصنع المعجزات:

إن التسويق يعتبر أفضل استثمار في العالم إذا قمت به على النحو الصحيح، وأداؤه على النحو الصحيح يتطلب تخطيطاً وصبراً وهو لا يصنع المعجزات. [8]

4.2.1.2 التسويق الذكي:

التسوق الذكي هو ما يحتاجه المستهلك اليوم، والتي حدد لها شرطاً أساسياً للحصول على جدواها، وهو عدم الاعتماد على الذاكرة في رصد الاحتياجات مباشرة من أمام رفوف مراكز التسويق . ومن

الضروري إعادة الأولويات والميزانيات مع أهمية إنشاء مراكز معلومات ومواقع لترشيد الاستهلاك ، وذلك يساهم في انضباط السلوك أثناء عملية التسوق وعدم الوقوع في مصيدة قرار الشراء السريع المبني على توقع مستوى إشباع لا يتطابق مع المستويات الفعلية التي تتحقق بعد البدء في استهلاك السلعة .

5.2.1.2 استراتجيات التخطيط للتسويق الذكي:

- رسالة الشركة : لابد من وجود رسالة واضحة للشركة ليعرف الناس عن شركتك.
- الاهداف الرئيسية : وجود اهداف رئيسية تساعد صاحب راس المال على التركيز في عمله وابتكار خطوات وطرق من اجل تحقيق هذه الاهداف الرئيسية.
- المراجعة التسويقية : مراجعة السوق والطريقة التسويقية للمنتجات او الخدمة التي يقدمها صاحب راس المال ، وهل هناك تقبل لطريقة التسويق التي يستخدمها صاحب المنتج او الخدمة .
- تحليل الشركة والبيئة والمستهلك :دراسة السوق والبيئة والمستهلك لتجنب اي خسائر او مشاكل قد تواجه الشركة مستقبلا.
- وضع الفرضيات لمواجهة اي تحدي : وضع حلول افتراضية مستقبلية لمواجهة اي مشاكل او تحديات لبعض المنافسين او مشاكل حروب او اضطرابات سياسية واخرى.
- الاهداف التسويقية والاستراتجيات : طريقة وضع الاهداف التسويقية والخطط التسويقية والاستراتجيات المستخدمة من فترة الي فترة معينة على شكل تسلسل زمني وتسلسل من الانجازات التي تخدم اهداف الشركة.
- تقدير النتائج المتوقعة : وضع نتائج متوقعة بين كل فترة زمنية معينة لقياس النتائج والمشاكل التي تواجه الشركة او المشروع.
- وضع الخطط البديلة والطوارىء :الخطط البديلة التي قد تطرأ بسبب الحروب او مشاكل سياسية او اقتصادية.
- الموازنة : وضع ميزانية مناسبة للتخطيط وهي المحرك الوحيد لكل الخطط وهي احدى الاسباب الرئيسية لنجاح المشاريع وهوالقلب الذي يبث الحياه في الشركة والمشاريع الخ. [9].

6.2.1.2 التسويق المصرفي:

فى البداية يحظى تسويق الخدمات المصرفية باهتمام كبير من جانب القائمين على إدارة البنوك التجارية، حيث تتنافس هذه البنوك فيما بينها سعياً وراء زيادة حجم عملياتها وذلك عن طريق جذب عملاء جدد، ومن هذا المنطلق قامت الكثير من البنوك بإنشاء إدارة مستقلة للتسويق المصرفي عهد إليها بمسؤولية وضع البرامج التي تمكن البنك من تحقيق هذا الهدف.

ماهية التسويق المصرفي:

1-مفهوم التسويق المصرفي:

مجموعة الأنشطة اللازمة لتسهيل عملية التبادل بين البنك والعميل بهدف إشباع رغبات العميل من جهة وتحقيق أهداف البنك من جهة أخرى، وعملية التسويق المصرفي تبدأ بالتعرف على رغبات العميل ثم تنتقل إلى تطوير الخدمات المناسبة التي تلبي تلك الرغبات ثم تحديد السعر المناسب والوقت المناسب لتقديم تلك الخدمات، بالإضافة إلى تحديد السياسات الترويحية اللازمة لتقديم هذه الخدمة، وعليه فإن التسويق المصرفي يبدأ وينتهي مع العميل بمعنى أن العميل (هو محور النشاط التسويقي للبنك)

إن التسويق المصرفي يجب أن يكون المبدأ الذي تدار بناء عليه المؤسسة المصرفية وليس مجرد نشاط تقوم به إدارة من إدارات تلك المؤسسة.

2- أهمية التسويق المصرفي(Importance of Banking (Marketing

يتكون النشاط التسويقي من مجموعة من الأنشطة التي تشمل عمليات التخطيط والتوجيه ورقابة الأداء بهدف تطوير الوضع التنافسي للبنك داخل السوق المصرفية، ويتم ذلك من خلال العمل على الحفاظ على العملاء الحاليين وتوسيع دائرة العملاء عن طريق الخدمات المصرفية بشكل متطور ومتميز، وقد ظهرت الحاجة لعملية التسويق نتيجة لازدياد حدة المنافسة بين البنوك التجارية من جهة والمؤسسات المالية الأخرى من جهة أخرى، بالإضافة إلى التغيرات السريعة في المؤثرات الاقتصادية والاجتماعية التي تحكم نشاط هذه المؤسسات والتي بدورها أوجبت على البنوك العمل على مواكبة هذه التغيرات للمحافظة على مركزها التنافسي.

3-أهداف التسويق المصرفي(Objectives of Banking Marketing):

يرتكز نشاط التسويق المصرفي حول تحقيق أهداف البنك الرئيسية والتي يأتي في مقدمتها زيادة ربحية البنك وقد تختلف تلك الأهداف من بنك لآخر حيث قد تضع بعض البنوك هدف زيادة حصتها السوقية في مقدمة الأهداف التي تسعى لتحقيقها، وأيا كان الهدف فإن مهمة التسويق المصرفي تركز على تطوير الوسائل اللازمة لتمكين البنك من تحقيق أهدافه، ومن أمثلة هذه الوسائل تطوير خدمات متميزة أو المساهمة في أنشطة اجتماعية معينة[18]

7.2.1.2 التسويق بإستخدام التنقيب في البيانات:

البيانات يمكن ان تعرف الى مجموعات حسب تركيبها بإعتبار المواضيع المختلفة مثال(بلدان , ولايات) بالملاحظة المتعددة ومثال على ذلك (سنوي, شهريا , فصلي او كل ساعة) ليتم تتبع تغيير المواضيع بمرور الوقت وفي نفس الوقت للتغلب على مشكلة الملاحظات المحدودة لكل موضوع Panel data analysis يستخدم كثيرا ي علوم الاجتماعيات مثل الاقتصاد.

الشركات التجارية الصغيرة التي تتمتع بإدارة جيدة تشكل بطبيعة الحال العلاقات مع الزبائن مع مرور الوقت يعرفون المزيد والمزيد عن عملائها , ويستخدمون هذه المعرفة لخدمتهم وهذا افضل.

8.2.1.2 تحليل وإدارة علاقات العملاء:

الشركات التجارية يجب ان تعمل علاقة واحد لواحد(one_to_one) مع عملائها.

أدارة علاقات العملاء هو موضوع واسع للعديد من الكتب والمؤتمرات يتم عمل برامج تعرف بالمركز للأدارة ويتم تسويق البرمجيات الآن كأداة لإدارة علاقات العملاء . [3]

2.2 الدراسات السابقة :

1.2.2 الدراسة الأولى: دور تنقيب البيانات في زيادة أداءالمنظمة(دراسة تحليلة في المصرف الصناعي ببغداد):

في هذه الدراسة تم اجراء عملية تنقيب البيانات على بيانات القروض للمصرف الزراعي العراقي بغرض مساعدة الادارة على إتخاذ القرار الذي يمكنها من الحكم على الزبون اذا كان التعامل معه يشكل خطرا على المصرف ام لا وبذلك نتمكن من التركيز على الزبائن الموثوق بهم واستبعاد غير الموثوق بهم .

تم إستخدام خوارزميتين اساسيتين لتحليل البيانات هما:-

1- شجرة القرار.
2- التجميع.

إن النماذج الناتجة من إستخدام خوارزميتي شجرة القرار والتجميع تلخص لإدارة المصرف بشكل واضح مجتمع البيانات وتسهل رسم الأستنتاجات منه فالخوارزمية الاولى بينت ان على الادارة ان تعيد النظر بالقروض ذات المستوى الادنى لان هذه القروض تشكل نسبة الخطر الاكبر كقروض غير مسددة اى ان على الادارة أن تكون اكثر حذرا من هؤلاء المقترضين وتحاول التقليل منهم وتركز على القروض ذات المستوى الأعلى .فيما صنفت الخوارزمية الثانية المقترضين في مجموعات وبشكل تنازلي بحيث يسهل التعرف على المجموعات بنظرة واحدة ومعرفة خصائصها.

مميزات الدراسة:

1- حلت الدراسة مشكلة تراكم البيانات.

2- تحسين مستوى إستخدام المعلومات من قبل الإدارة من اجل زيادة الأداة .

3- مساعدة المصرف علي تجنب مخاطر حقيقية ناتجة من عدم التسديد.

2.2.2 الدراسة الثانية: دور تقنية تنقيب البيانات في التسويق وإدارة العلاقة مع الزبائن بالنسبة لاعمال الاتصالات(حالة دراسية : شركة جوال):

هذه الدراسة تركز على دور تقنية تنقيب البيانات في دعم إستراتيجية التسويق وإدارة العلاقة مع الزبائن بالنسبة لاعمال الاتصالات وهنالك عدد من المحاور التي تركز عليها الدراسة وهي الاحتياجات الالكترونية اللازمة في نظام إدارة العلاقة مع الزبائن لبناء تطبيقات تنقيب البيانات ,التطبيقات والانشطة التسويقية المرتبطة بتقنية تنقيب البيانات من ناحية تحليل خصائص الزبائن و تحديد الجهة المستهدفة ,استقطاب زبائن جدد والتسويق المباشر .

مميزات الدراسة :

فقد اثبتت الدراسة ان هذه التقنية تستخدم بشكل فعال في تتبع سلوك الزباين والتنبؤ بالزباين الذين من المحتمل ان تفقدهم الشركة والزبائن الاكثر ربحية.

3.2.2 الدراسة الثالثة :(تسويق المحتوى من خلال استخراج البيانات من الشبكة الاجتماعية الفيسبوك):

في هذه الدراسة، يقترح تسويق المحتوى كأسلوب للتسويق في الشبكات الاجتماعية .في المرحلة الأولى يتم دراسة معلومات حول سلوك مستخدمي الفيسبوك ويتم اختيار عدد من المستخدمين بناء علي سلوكهم في الماضي وفعالية مواقعهم بين المستخدمين الاخرين من اجل توزيع المعلومات من خلالهم .

في المرحلة الثانية من خلال الإستفادة من تقنيات استخراج البيانات , وتحليل سلوك المستخدمين الفردي والجماعي وفقا لتفاعل المستخدمين في منشوراتهم وبناء علي ذلك يتم تزويدهم بمحتوياتهم الخاصة .

يتم انشاء حساب الفيسبوك وتحليل سلوك المستخدمين ويستخدم التطبيق Netvizzهذا التطبيق يتيح للباحثين استخراج البيانات المطلوبة من مختلف الخدمات من الفيسبوك وحفظ أو عرض ملفات الإخراج في اشكال مختلفة.

مميزات الدراسة :

1- في هذا النوع من التسويق لا يشترط مقابلة المستخدمين وجها لوجه.

2- الشبكة الاجتماعية الفيسبوك هي أساسا مناسبا لتسويق المنتجات , حيث منتجات مختلفة يمكن تقديمها بسهولة وتوزيعها بين المستخدمين بسرعة.

3- يتم خلالها تطوير وتعزيز المعارف والمعلومات بين المستخدمين، والمحتوى التسويقي يخلق شعورا من الثقة في نفوسهم لشراء السلع من خلال المحتوى المقدم .

جدول (1.2)مقارنة بين هذه الدراسة وتادراسات السابقة

رقم الدراسة	وجه الشبه	وجه الإختلاف
1	إستخدمت الدراسة تقنية التنقيب في البيانات وخوارزميات شجرة القرار	إستخدمت خوارزميات التجميع
2	إستخدمت تقنيات في البيانات في تتبع سلوك الزبائن	لم تستخدم خوارزمية شجرة القرار
3	إستخدمت تقنيات التنقيب في البيانات في تحليل سلوك المستخدمين في الفيسبوك	إستخدمت تطبيق Netvizz لإستخراج البيانات من الفيسبوك

الباب الثالث

التطبيق

3.1. مقدمة عن التطبيق:-

في هذا المشروع استخدمنا كمية كبيرة من البيانات عن عملاء للبنك ثم قمنا بتطبيق كل خطوات التنقيب عن البيانات حتى نقوم بعمل نموذج يساعد البنك في تصنيف العملاء لمعرفة اذا كان العميل سوف يودع وديعة ام لا واستخدمنا لذلك خوارزميات شجرة القرار لانها تتميز بسهولة فهم النتائج .

3.2. التقنيات المستخدمة:

3.2.1. برنامج(SPSS Clementine):

وهو واحد من أول أدوات جمع البيانات للأغراض العامة وهو الاكثر إستخداما في حزم إستخراج البيانات .

وهو عبارة عن تقنية تسمح للخبراء بتنقيب البيانات الخاصة بهم ولديه واجهة تدفق للبيانات تقوم بتبسيط عملية التنقيب في البيانات.

تشمل التطبيقات(SPSS Clementine)تقسيم العملاء والتنميط لشركات التسويق، والكشف عن الغش، وسجل الائتمان، والتنبؤ بالربح لتجار التجزئة.[10]

حقائق حولSPSS Clementine:-

1- يدعم منهجية(CRISP-DM) لأستخراج البيانات.
2- لا يقتصر على خوارزمية تنقيب عن البيانات واحدة .أنما يحتوي على مجموعة من الخوارزميات .
3- ليس من الضروري ان يعرف المستخدم كيفية عمل الخوارزمية.[4]

3.2.2. HTML:

هي اللغة الترميزية القياسية المستخدمة لإنشاء صفحات الويب. يمكن لمتصفحات الويب قراءة ملفات (HTML) وجعلها في صفحات الويب مرئية أو مسموعة. عناصر (HTML) تشكل اللبنات الأساسية لجميع المواقع. (HTML) يسمح الصور والكائنات لتكون جزءا لا يتجزأ، ويمكن استخدامها لإنشاء نماذج تفاعلية. فإنه يوفر وسيلة لخلق وثائق المنظمة التي تدل دلالات هيكلية للنص مثل العناوين والفقرات والقوائم وصلات.[13][14]

من مزايا html:-

1-من السهل فهمها واستخدامها.

2-هي لغة متاحة مجانياً .

3- تدعمها جميع المتصفحات.[15]

3.2.3. php :-

هي لغة البرمجة النصية من جانب الخادم المصممة لتطوير الشبكة ولكنها تستخدم أيضا كلغة برمجة للأغراض العامة[16]

من مزايا php:

1-سهولة التعلم والقدرة العالية.
2- السرعة العالية في تنفيذ البرامج.
3- توافقها مع جميع أنظمة التشغيل , ودعمها لعدد كبير من قواعد البيانات .
4- تحتوي على كثير من الدوال الجاهزة للاستخدام .[17]

3.2.4. MySQL:

هو نظام إدارة قواعد البيانات العلائقية مفتوح المصدر ,ويعمل على جميع أنظمة التشغيل ويندوز,لينكس,يونيكس, وهو نظام عالي الأداء ، و أصبح المعيار في إنشاء تطبيقات قواعد البيانات على الويب أو خارجها.[11]

من مزايا MySQL:-

1-سهلة التعلم .

2- مفتوحة المصدر.

3- غير مكلفة.

3- تدعم بسهولة كلما دعت الحاجة لها.[12]

3.3. وصف البيانات:

تم الحصول على البيانات من الموقع الالكتروني **(Uci machine leaning repository)** وذلك لعدم امكانية الحصول على بيانات من البنوك نظرا لحرصهم على سرية بيانات العملاء فيها .

هذا وصف للبيانات التى تم استخدامها:

عدد سجلات البيانات 41188 .

عدد الحقول 21 وهي:

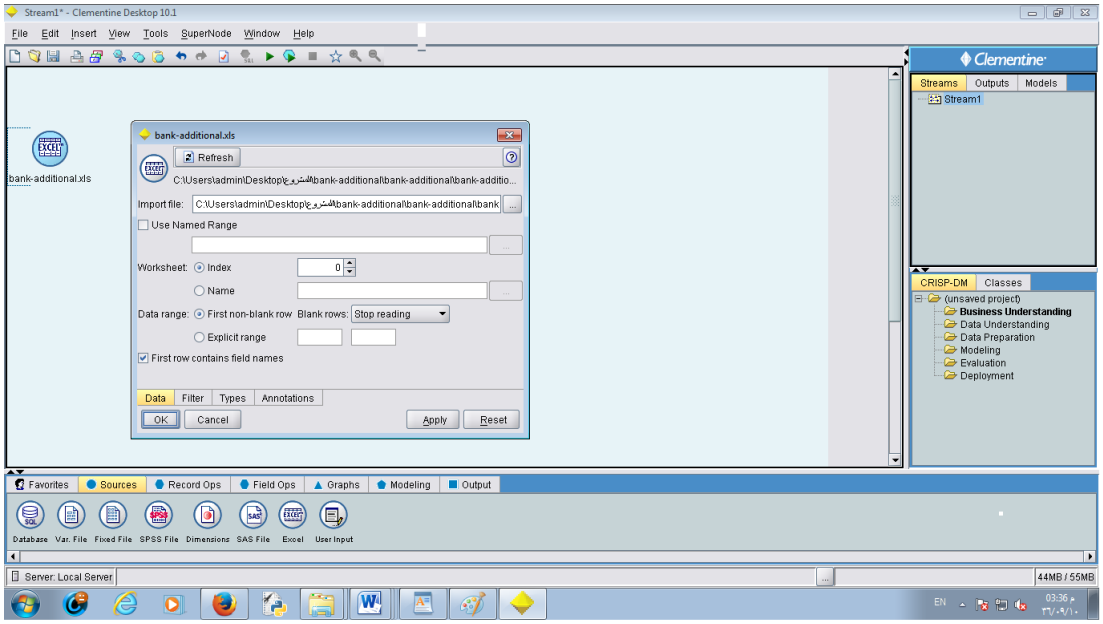
1 - age (رقم)
2 - job : نوع الوظيفة (categorical: "admin.","blue-collar","entrepreneur","housemaid","management","retired" "self employed","services","student","technician","unemployed" ,"unknown")
3 - marital : الحالة الاجتماعية (categorical: "divorced","married","single","unknown"; note: "divorced" means divorced or widowed)

4 - education المستوى التعليمي(categorical: "basic.4y","basic.6y","basic.9y","high.school","illiterate","professional.course", "university.degree","unknown")
5 - default: هل لديه ائتمان ؟ (categorical: "no","yes","unknown")
6 - housing: لديه قروض الإسكان؟ (categorical: "no","yes","unknown")
7 - loan: لديه قرض شخصي؟ (categorical: "no","yes","unknown")
#ذات صلة معالاتصالالأخير منالحملة الحالية :
8 - contact: نوع الاتصال (categorical: "cellular","telephone")
9 - month: اخر شهر تم فيه الاتصال (categorical: "jan", "feb", "mar", ..., "nov", "dec")
10 - day_of_week: اخر يوم تم فيه الاتصال (categorical: "mon","tue","wed","thu","fri")
11 - duration: المدة الزمنية للاتصال بالثانية (numeric). # other attributes:
12 - campaign: عدد الاتصالات للعميل خلال الحملة (numeric, includes last contact)
13 - pdays: عدد الايام التي مرت بعد الاتصال بالعميل في الحملة التسويقية السابقة (numeric) (999 تعني لم يتصل العميل سابقا).
14 - previous: عدد الاتصالات التي اجريت قبل هذه الحملة لهذا العميل (numeric)
15 - poutcome: نتائج الحملة التسويقية السابقة
(categorical: "failure","nonexistent","success")
سماتالسياقالاجتماعيةوالاقتصادية:
16 - emp.var.rate: معدل تغيير الموظفين - quarterly indicator (numeric)
17 - cons.price.idx: الرقم القياسي لأسعارالمستهلك - monthly indicator (numeric)
18 - cons.conf.idx: مؤشر ثقة المستهلك - monthly indicator (numeric)
19 - euribor3m: euribor 3 month rate - daily indicator (numeric)
20 - nr.employed: عدد الموظفين - quarterly indicator (numeric)

Output variable (desired target):
21 - y - has the client subscribed a term deposit? (binary: "yes","no).

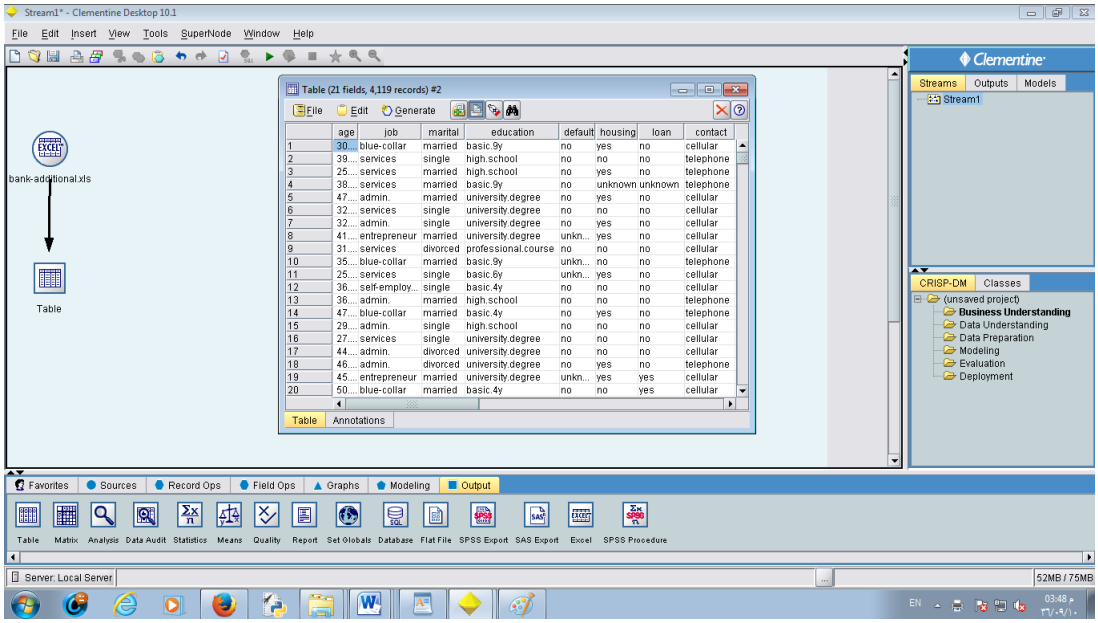
3.4. التطبيق على البيانات:

تم تحميل البيانات على برنامج (Clementine)من ملف (excel)عن طريق (node) الموضحة ادناه:



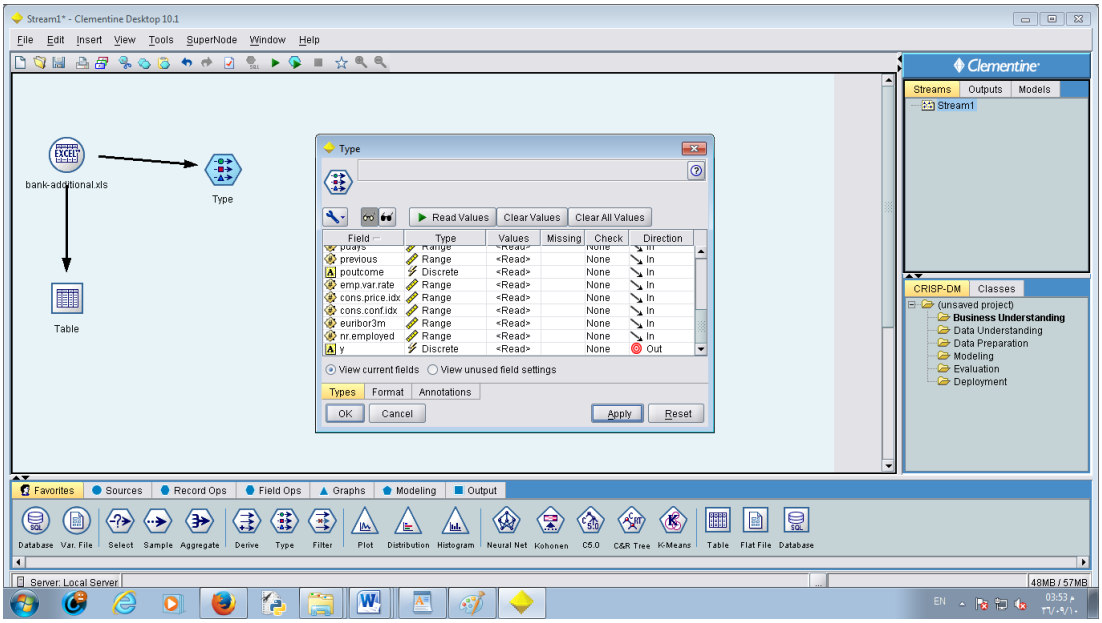
شكل(3.1) تحميل البيانات على البرنامج

إختيرت (node) ال (table) من (palettes output) بالضغط على (execute) يتم تحميل البيانات في الجدول كما هو موضح ادناه



شكل(2.3) تحميل البيانات في جدول

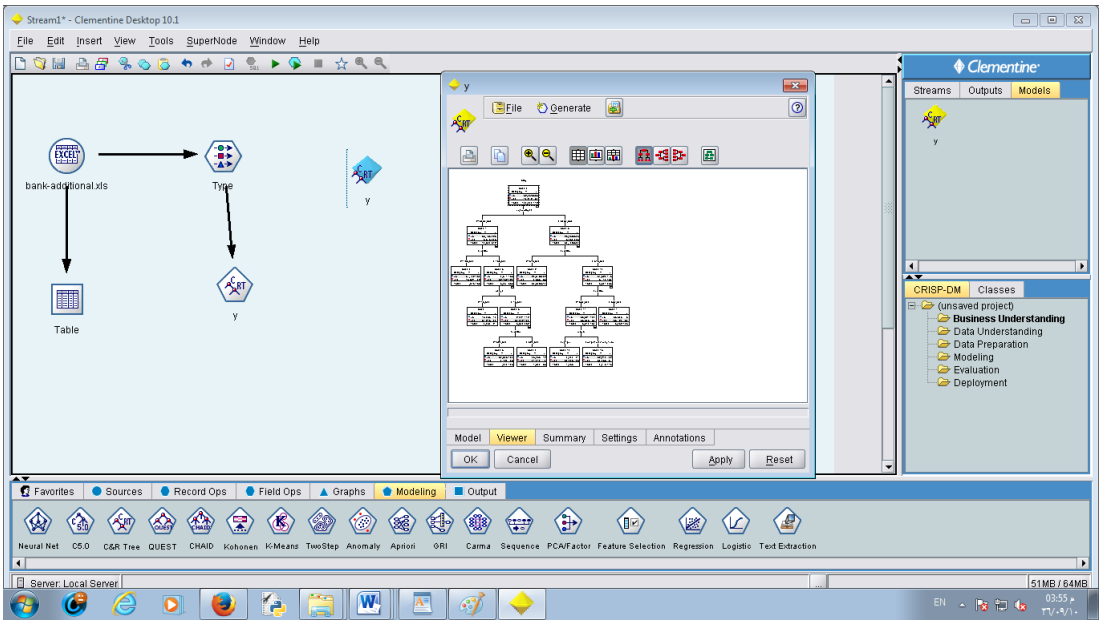
إختيرت(node) ال (type) من (palettes field operation) وذلك لتحديد الخاصية التي يعتمد عليها التصنيف وهي y كما في الشكل :



شكل(3-3) تصنيف البيانات على حسب y

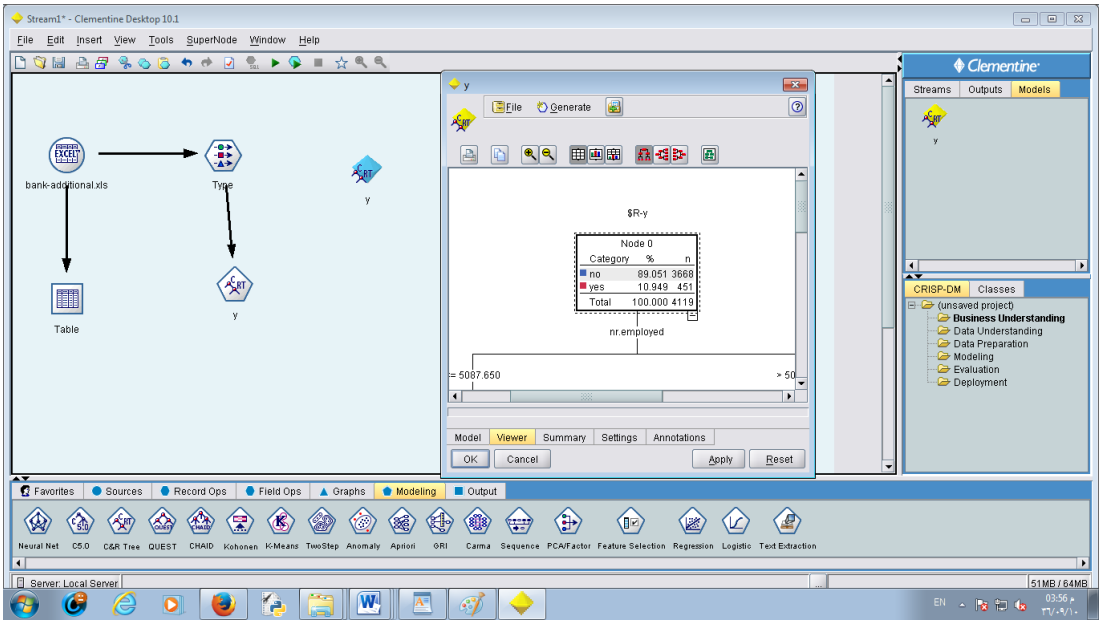
إختيرت (node) ال (c&r tree) من (palettes modeling)

(C&r tree) هي خوارزمية من خوارزميات شجرة القرار بعد ذلك يتم الحصول على نموزج باللون الاصفر يحتوي على نتائج التصنيف كالاتي:

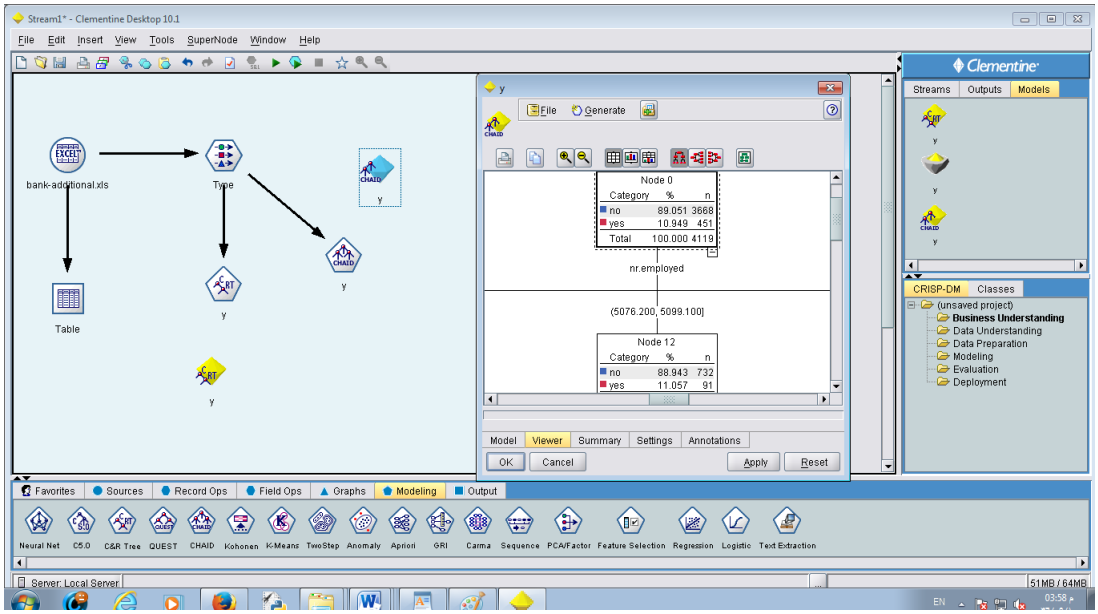


شكل(3.4) إختيار الخوارزمية المستخدمة

إختيرت خوارزمية اخرى وتم مقارنة نتائج الخوارزميتين للتأكد من صحة النتائج .

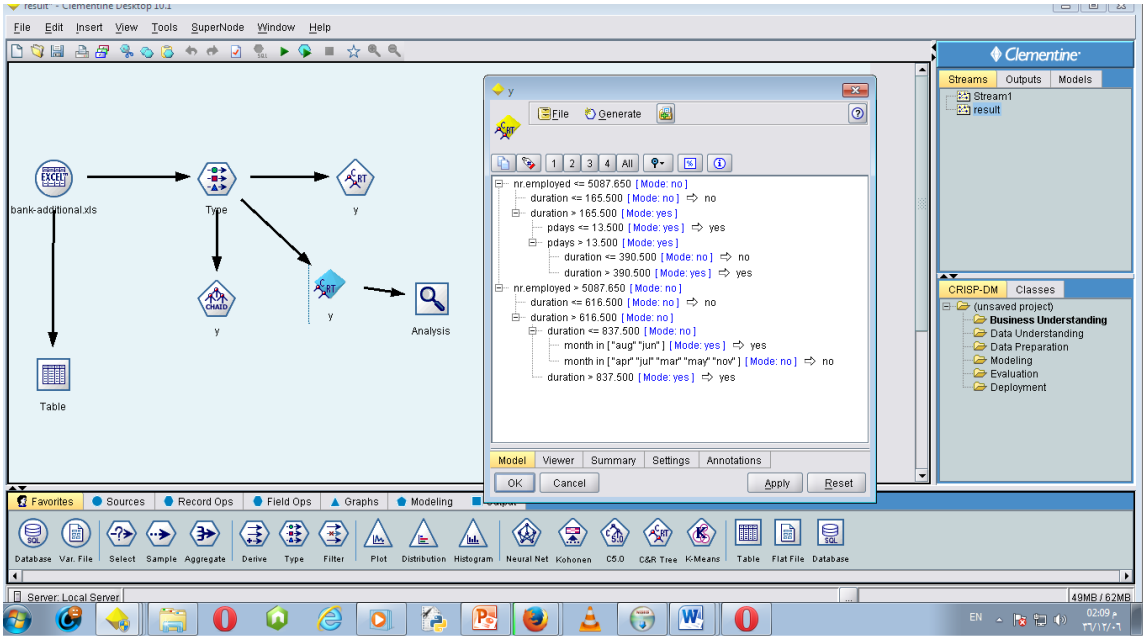


شكل(3-5) نتيجة الخوارزمية المستخدمة



شكل(3.6) خوارزمية اخرى

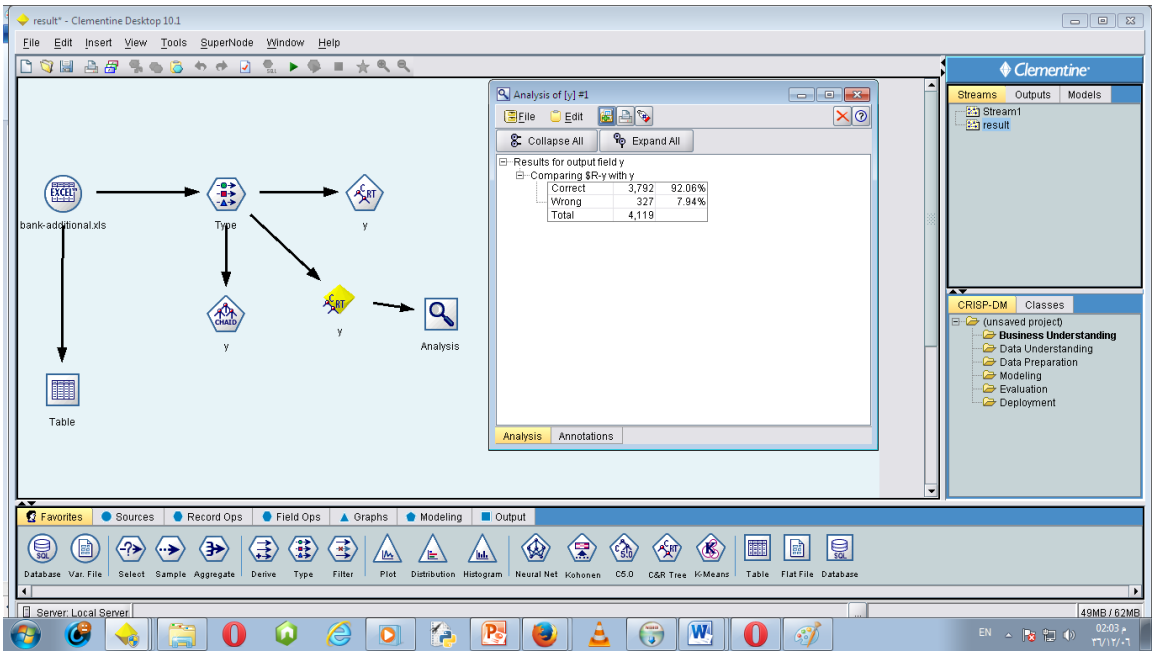
لقد تم الحصول على النتائج التالية من تحليل البيانات من خلال خوارزمية (C&R tree):-



شكل(7-3) نتائج تحليل البيانات الناتجة من الخوارزمية

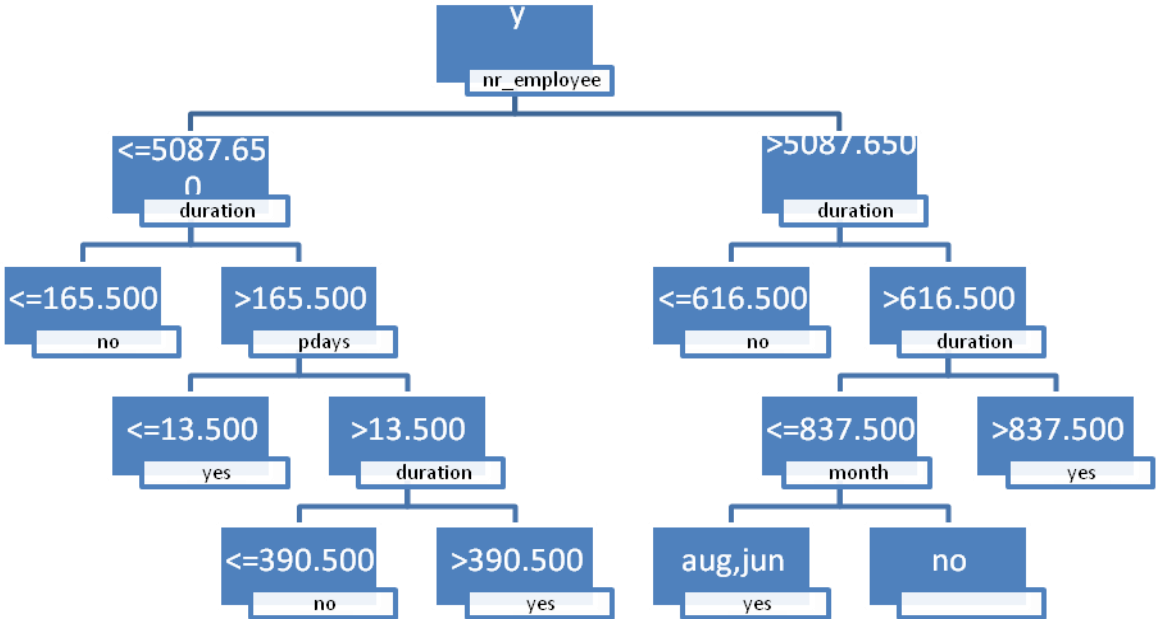
دقة الخوارزمية المستخدمة :-

دقة هذه الخوارزمية هي 92% وهذا يدل على تحليل البيانات بالصورة المطلوبة.



شكل(3.8) دقة الخوارزمية المستخدمة

النموذج الذي تم الحصول عليه للبيانات المخرجة :



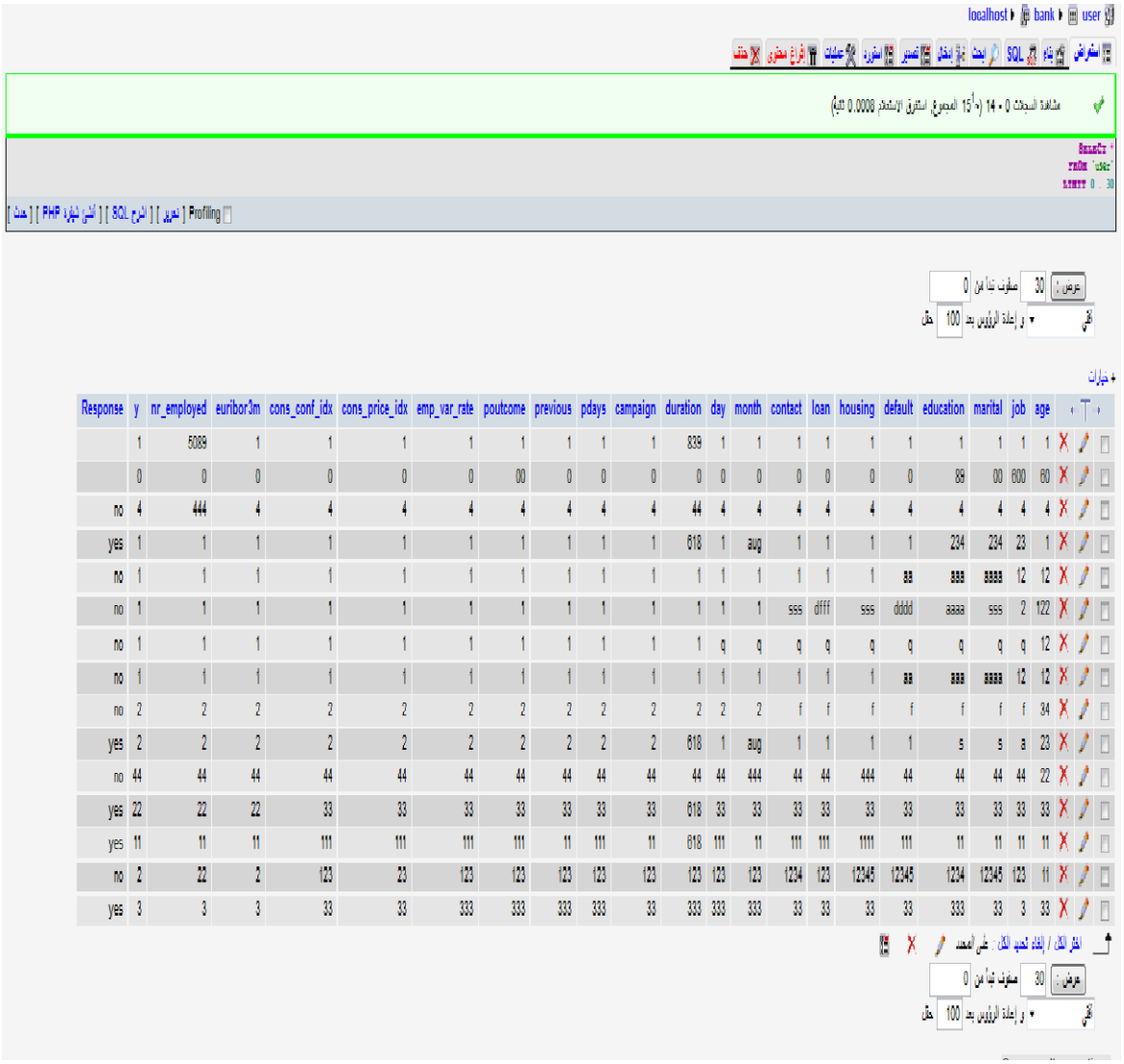
شكل(3.9) نموذج البيانات المخرجة

3.5. بناء واجهات النموذج:-

بعد أن تم الحصول على قواعد من تحليل البيانات أستخدمت هذه القواعد لبناء نموذج له واجهات إستخدام لكي يتم التعامل معه من قبل موظفي البنك ,استخدمت لغة php و html لعمل الواجهات .

3.5.1. تطبيق قاعدة البيانات :-

جدول بيانات العملاء:-



age	job	marital	education	default	housing	loan	contact	month	day	duration	campaign	pdays	previous	poutcome	emp_var_rate	cons_price_idx	cons_conf_idx	euribor3m	nr_employed	y	Response
1	1	1	1	1	1	1	1	1	1	839	1	1	1	1	1	1	1	1	5089	1	
60	800	00	88	0	0	0	0	0	0	0	0	0	0	00	0	0	0	0	0	0	
4	4	4	4	4	4	4	4	4	4	44	4	4	4	4	4	4	4	4	444	4	no
1	23	234	1	1	1	1	1	aug	1	618	1	1	1	1	1	1	1	1	1	1	yes
12	12	8888	888	88	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	no
122	2	sss	aaaa	dddd	sss	dfff	sss	1	1	1	1	1	1	1	1	1	1	1	1	1	no
12	q	q	q	q	q	q	q	q	q	1	1	1	1	1	1	1	1	1	1	1	no
12	12	8888	888	88	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	no
34	f	f	f	f	f	f	f	2	2	2	2	2	2	2	2	2	2	2	2	2	no
23	a	s	s	1	1	1	1	aug	1	618	2	1	1	2	2	2	2	2	2	2	yes
22	44	44	44	44	444	44	44	444	44	44	44	44	44	44	44	44	44	44	44	44	no
33	33	33	33	33	33	33	33	33	33	618	33	33	33	33	33	33	33	22	22	22	yes
11	11	11	11	111	1111	111	111	11	111	618	11	111	11	111	111	111	111	11	11	11	yes
11	123	12345	1234	12345	12345	123	1234	123	123	123	123	123	123	123	123	23	123	2	22	2	no
33	3	33	333	33	33	33	33	333	333	333	33	333	333	333	333	33	33	3	3	3	yes

شكل(3.10) جدول بيانات العملاء في قاعدة البيانات

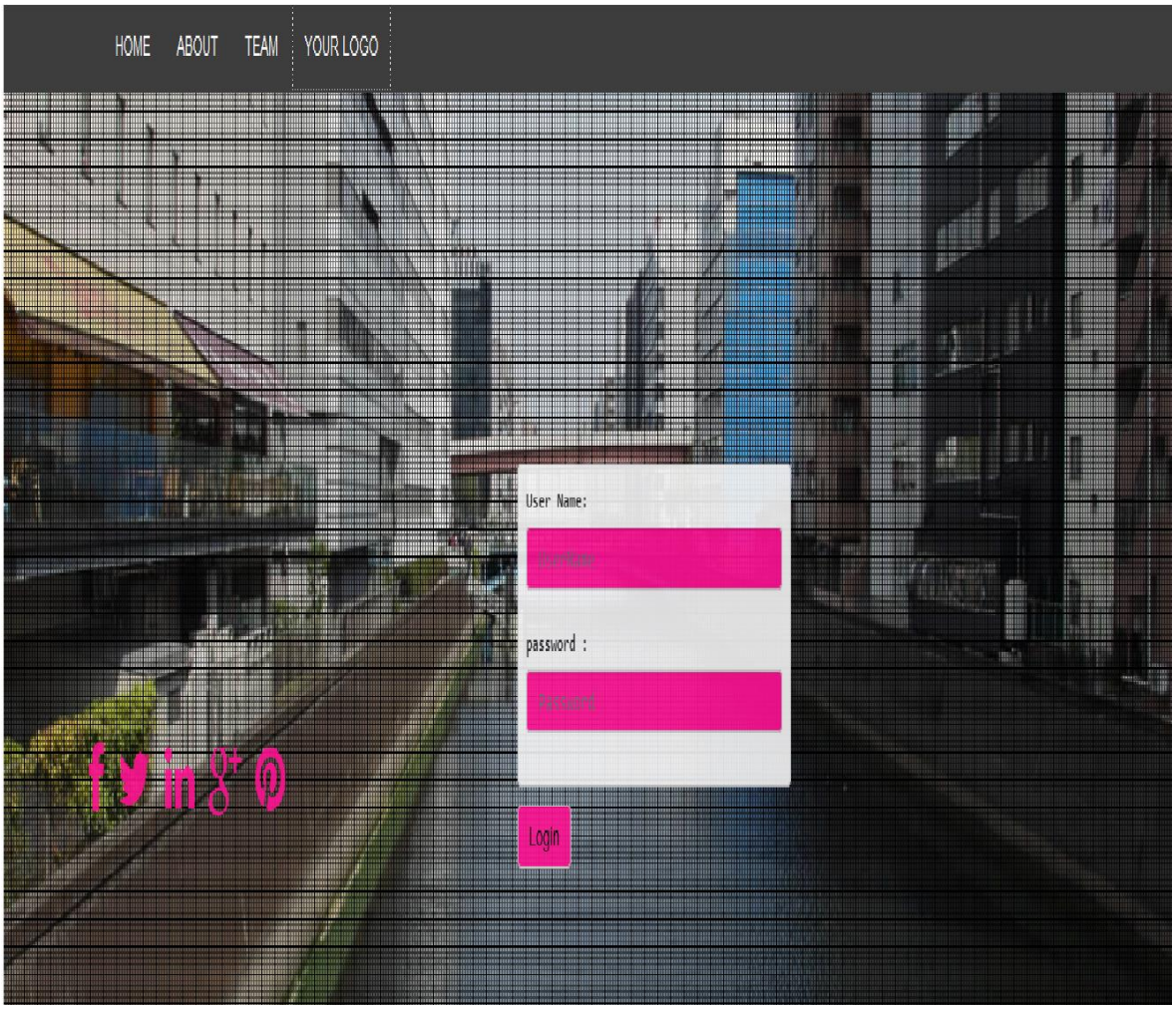
3.5.2. الواجهات:

1- الواجهة الرئيسية:-

شكل(3.11) الواجهة الرئيسية.

2-واجهة تسجيل الدخول:-

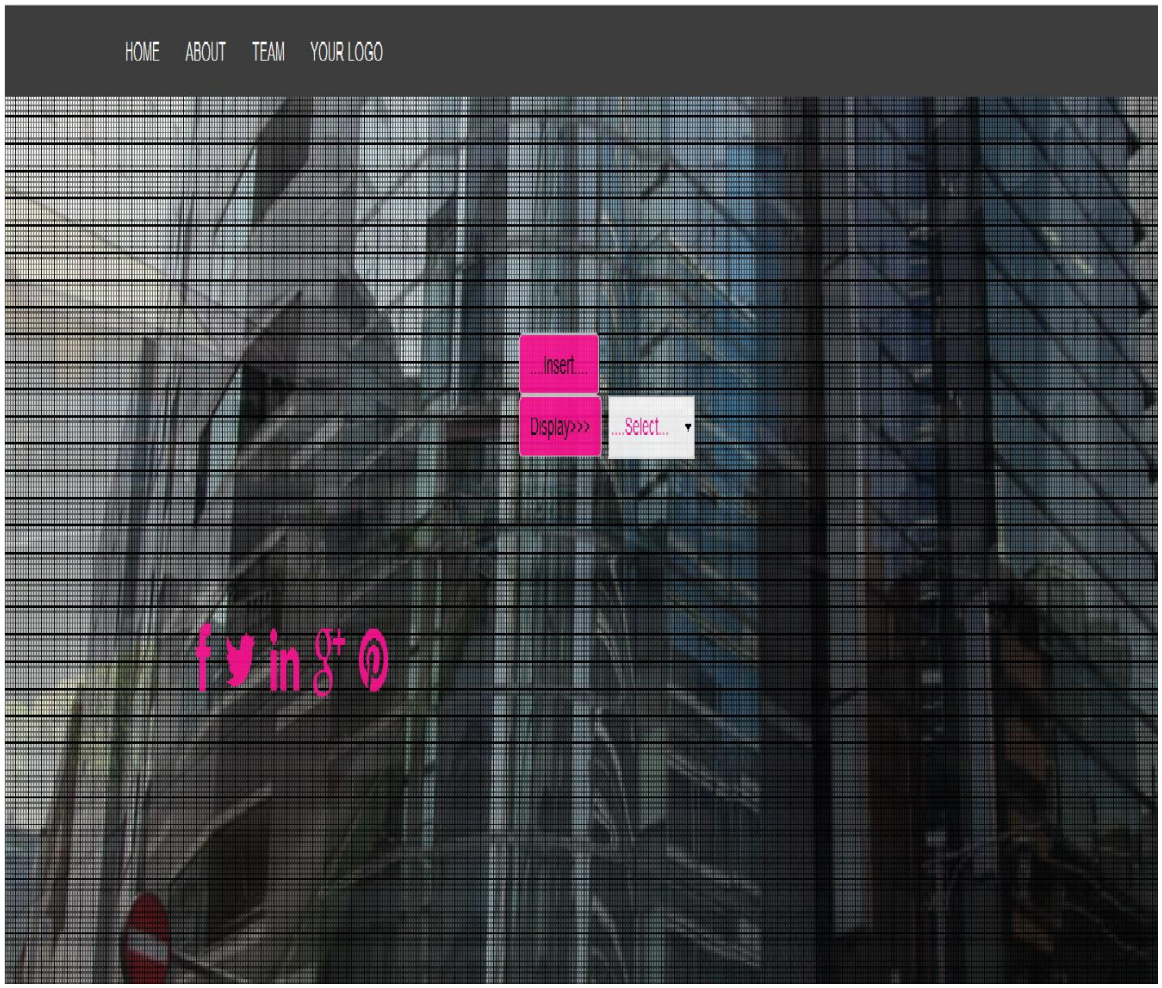
تسجيل دخول الموظفين بالبنك بعد كتابة إسم المستخدم وكلمة المرور



شكل(3.12) واجهة تسجيل الدخول.

3-واجهة الأختيار:-

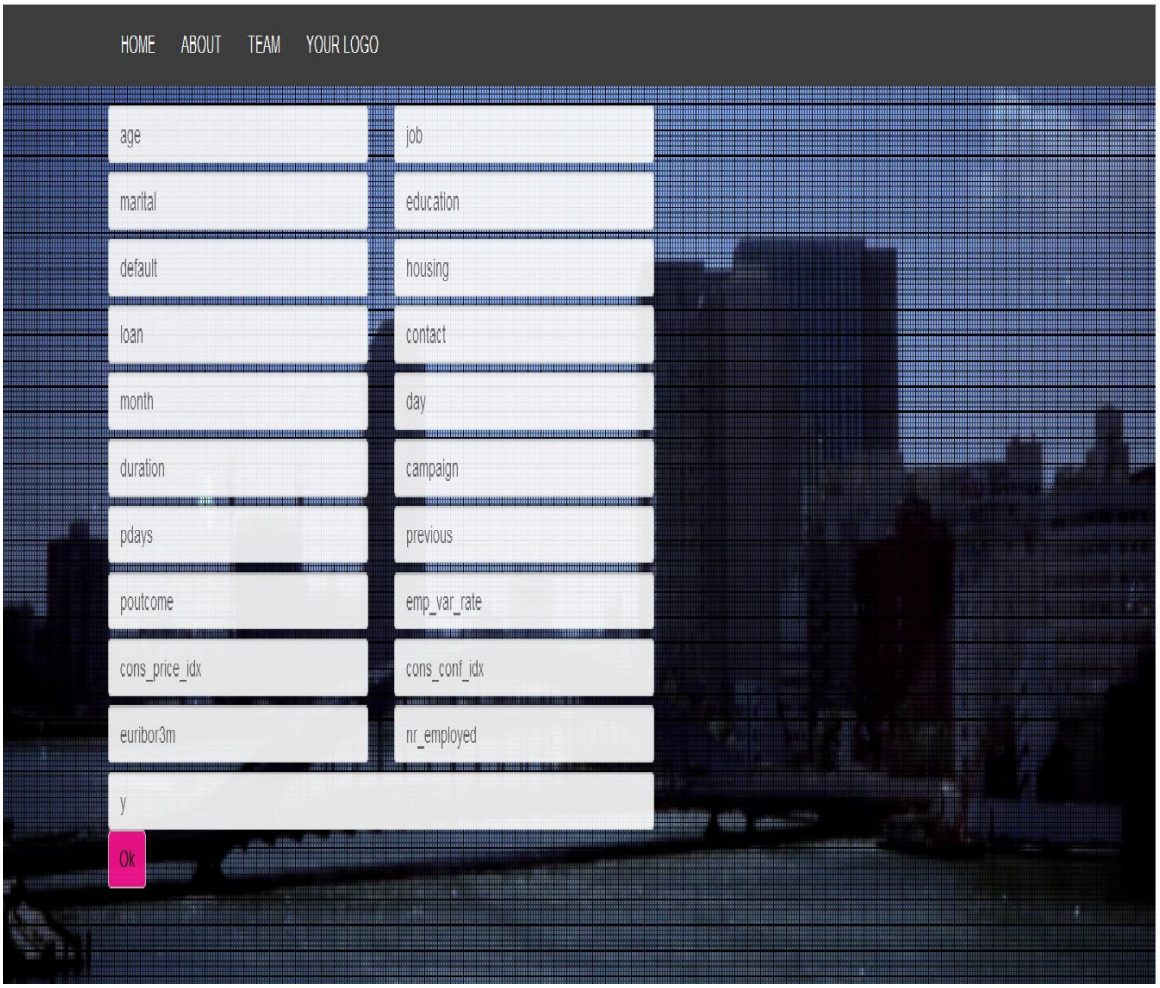
لكي يختار الموظف بين عملية إضافة عميل جديد أو عرض بيانات العملاء بناء على إستجابة العميل للحملة التسويقية.



شكل(3.13) واجهة الإختيار.

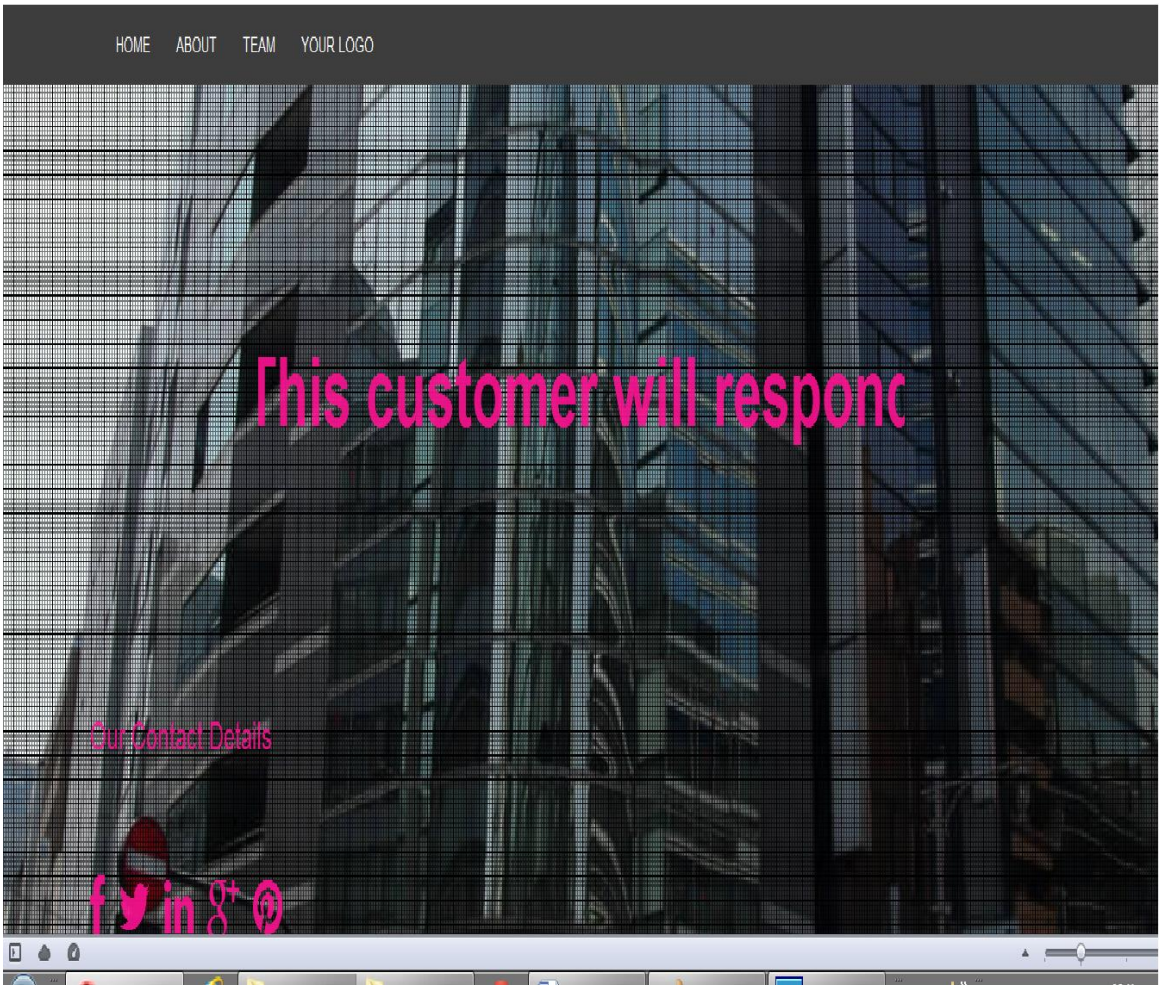
4-واجهة إضافة العملاء:-

تتم إضافة العملاء بإدخال جميع البيانات الخاصة بهم لمعرفة ما إذا كان العميل سوف يستجيب للحملة التسويقية ام لا .



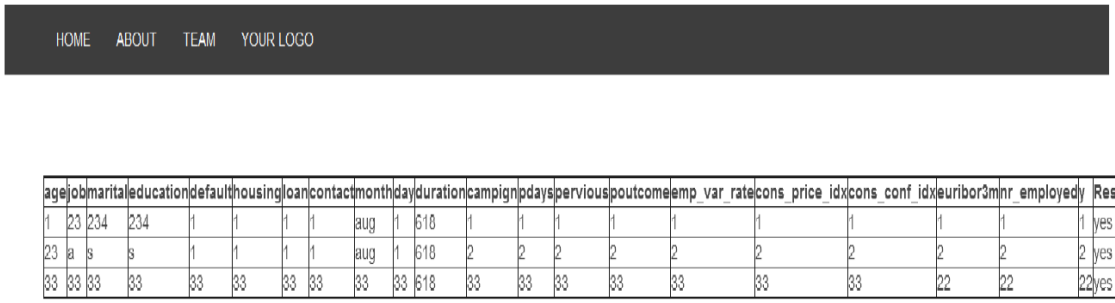
شكل(3.14) واجهة إضافة العملاء.

5- واجهة لمعرفة إستجابة العميل :



شكل(3.15) واجهة معرفة إستجابة العميل.

6- واجهة لعرض بيانات العملاء :



age	job	marital	education	default	housing	loan	contact	month	day	duration	campign	pdays	pervious	poutcome	emp_var_rate	cons_price_idx	cons_conf_idx	euribor3m	nr_employed	y	Response
1	23	234	234	1	1	1	1	aug	1	618	1	1	1	1	1	1	1	1	1	1	yes
23	a	s	s	1	1	1	1	aug	1	618	2	2	2	2	2	2	2	2	2	2	yes
33	33	33	33	33	33	33	33	33	33	618	33	33	33	33	33	33	33	22	22	22	yes

شكل(3.16) واجهة عرض بيانات العملاء.

الباب الرابع

النتائج والتوصيات والخاتمة

1.4 النتائج:

1. تم تصنيف العملاء حسب إستجابتهم للحملات التسويقية الى نوعين ,النوع الأول يستجيب للحملة التسويقية ويمثل 89.051% من العملاء, والنوع الثاني لا يستجيب للحملة التسويقية ويمثل 10.949% من العملاء.
2. نجاح خوارزمية شجرة القرار في تصنيف بيانات العملاء بدقة (92.06%) مما يدل على تحليل البيانات بالصورة المطلوبة .
3. بعض العملاء صنفوا على إنهم سوف يستجيبوا للحملة التسويقية ولكن في الواقع لم يستجيبوا ,هذا النوع من العملاء يجب على الإدارة التعامل معهم بصورة ذكية.
4. العينة غير متجانسة أى أن نسبة التصنيف غير متساوية , وهذا يحدث لكل نماذج تعلم الآلة , ولكننا في هذا البحث إستخدمنا البيانات الموجودة في الموقع لأننا لم نتمكن من الحصول عليها من البنوك السودانية .
5. حفظ زمن العملاء بإطلاعهم على الاعلانات المفيدة لهم فقط , وجعل التسويق فعال في البنوك.
6. خوارزمية شجرة القرار سهلة الفهم والإستخدام , والقواعد الناتجة منها واضحة .

2.4 التوصيات:

1. عمل مصنف لبيانات عملاء شركات الإتصالات لجعل التسويق فعال وأكثر كفاءة وحفظ موارد المستخدم .

2. يجب تخزين البيانات بصورة جيدة حتى تكون في متناول اليد لإكمال مسيرة الأبحاث في هذا المجال.

3. إستخدام تقنيات أخرى (weka).

4. نظراً لأهمية علم التنقيب في البيانات يجب تدريسه في الكلية بصورة أوسع.

الخاتمة

وبحمد الباري ونعمة منه وفضل ورحمة

وفي النهاية لانملك إلا أن نقول أننا قد أدلينا بفكرتنا في هذا الموضوع لعلنا نكون قد وفقنا في كتابته والتعبير عنه وأخيرا ما نحن إلا بشر قد نخطئ ونصيب فإن أخطأنا فلنا شرف المحاولة والتعلم وإن أصبنا فذاك مرادنا آملين أن ينال القبول والإستحسان

وصل الله على سيدنا محمد وعلى آله وصحبه وسلم.

المراجع

1. Radhakrishnan , Shineraj , anver Mohammed K.M, application of data mining in marketing (IJCSN International Journal of Computer Science and network , volume 2,issue 5 ,2013).
2. jiawe han, micheline kamber , jian pei, data mining :concepts and techniques ,3rd ed.(morgan kaufmann publisher,2006).
3. katsutoshi yada , data mining for marketing (faculty of commerce sansei university ,2011).
4. Dave Norris, Clementine data mining workbench from SPSS (Bloor Research 2005).
5. jiawe han, micheline kamber , jian pei, data mining :concepts and techniques ,2rd ed.(morgan kaufmann publisher,2006).
6. هالة الطويل , المرجع التعليمي في التنقيب عن البيانات (شعاع للنشر والعلوم,2009).
7. د.علاء حسين الحمامي , تنقيب البيانات(إثراء للنشر و التوزيع,2008).
8. .مجدي محمد عبد الله التسويق الذكي (ماهي للنشر والتوزيع وخدمات الكمبيوتر,2009).
9. عروة العريفي, إستراتيجيات التخطيط للتسويق الذكي (2014).
10. http://www.the-data-mine.com/Software/SPSSClementine

    التاريخ 2015/9/1 الزمن 10:00 ص .
11. http://searchenterpriselinux.techtarget.com/definition/MySQL

    التاريخ 2015/9/10 الزمن 1:20م.
12. https://www.datarealm.com/blog/five-advantages-disadvantages-of-mysql /

    التاريخ 2015/9/11 الزمن 10:30م.
13. Hypertext Markup Language | Definition of hypertext markup language by Merriam-Webster

    التاريخ 2015/9/11 الزمن 11:00م.

14. "HTML 4.0 Specification — W3C Recommendation — Conformance: requirements and recommendations". World Wide Web Consortium. December 18, 1997. Retrieved July 6, 2015.

التاريخ 2015/9/11 الزمن 11:30م.

15. http://www.rapidprogramming.com/questions-answers/What-are-the-advantages-of-HTML--508

التاريخ 2015/9/11 الزمن 11:50م.

16. Ide, Andy (2013-01-31). "PHP just grows & grows". Retrieved 2013-04-01

التاريخ 2015/9/11 الزمن 12:30م.

17. /http://www.sakshay.in/blog/10-advantages-of-choosing-php-web-development

التاريخ 2015/9/11 الزمن 1:00م.

18. http://www.startimes.com/f.aspx?t=31512474

التاريخ 2015/5/11 الزمن 1:00م.