

DATA MINING AND NEURAL NETWORKS II DMX USE FOR RISK ASSESSMENT OF COMPLICATIONS OF ARTERIAL HIGH BLOOD PRESSURE

MBUYI MUKENDI Eugène¹, KAFUNDA KATALAYI Pierre², MBELU MUTOBABevi³

¹Professor of the University of Kinshasa, Department of Computer Sciences, DR Congo

² University of Kinshasa, Department of Computer Sciences, DR Congo

³University of Kinshasa, Department of Computer Sciences, DR Congo

Kinshasa Computer science Laboratory.

Abstract

The results obtained in our paper on data mining and neural networks II provide probabilities of a complication when one factor is present [8]. Because of this, to assess the risk that one patient presents is to show one of the three kinds of complications, that is to say cerebral vascular accident, acute renal insufficiency and different heart disorders, according to the value taken by different factors simultaneously. We are going to query the model of data mining. The Language Data Mining Extensions (DMX) is used to create the structure of new models of Data Mining for learning, in other words, models must do prediction. [5]

Keywords: Data Mining Extension, DMX request, chart, Data Mining.

I. INTRODUCTION [5], [6], [8]

We must design a system enabling the assessment of risk that presents a patient of high blood pressure to develop one of the three types of complications of this one. The system should allow insertion of any case of new complication in the database, update the data mining structure and the data models which will consist for them in their learning when the database is modified.

For this assessment, we implemented software in Microsoft Visual Studio 2008, able to retrieve values of various factors, and on basis of this, to assess the risk for each of the three types of complications.

The system will have as end-user any member of the medical profession can get access. The headmaster of the system will have the possibility to use all functionalities of the system including those intended to the keying of data mining which will not be accessible to any end-user.

The system has to be designed in a way to support a future expansion on the network in order to reach a greater use in the area.

II. MODELING [5], [6], [11]

To model the system: applied principles of modeling in UML 2.0. This choice was motivated by the fact that we have undertaken to address the implementation of the system according to object-oriented approach which offers more advantages.

II.1. Diagram of class

From the analysis of initial specifications and knowledge on the field, we have identified classes hereafter:

- Server Analysis Services
- Structure of data mining
- Model of data mining
- Case of complication of high blood pressure
- risks of complications for patients
- Patients
- End-users
- System webmaster
- Factors

classes related to data processed. Therefore, it contains Patients, Factors, Risk and case Complication classes.

Associations related to different classes are as follows:

- The server hosts package of database, mining structures and data mining models
- Models learn a structure
- Mining structures restrict the database package
- The database webmaster updates mining structures and data mining models. He manages data and database users
- The end-users handles database
- Patient develops a case of complication get complications
- Factors allow evaluation of risk of complication
- Patients present risks of complication

II.2. Diagram of state

Classes presenting states in succession of our analysis are the following:

- Analysis server
- Mining Structure
- Mining Model

II.3. Use-case diagram

The main functions described in the original specifications are:

- Registration of new complication cases in the database
- Updating the data mining models
- Updating data mining structures
- Assessment of complication risk patients

Main users of the system are:

- End-users who is a member of the medical staff
- The webmaster of the system

II.4. Diagram of sequence

We have deduced only one sequence diagram which for us, is the most relevant because it refers to the use-case "updating data mining model" The title cannot determine all actions that must take place there. This

diagram describes an interaction between the webmaster and the Analysis Services server.

II.5. Activity diagram

As well for diagram of sequence seen above, we have described the activity of the use case "updating data mining model".

III. Notion on DMX requests [5], [6], [8]

During the creation of our data mining model, we used Microsoft Neural Network algorithm to perform the process of data mining on our database. Learning of this model allows only to get the importance of factors in observing a complication of high blood pressure. Our aim is to use this model for prediction (predict the risk of complications). To do so, we used DMX.

DMX (Data Mining Extensions) is a language close to SQL, introduced by Microsoft for handling data structures and data mining models created with Microsoft SQL Server from version 2000.

The purpose of DMX is to provide common concepts and common query expressions related to the data mining area, independently on the used algorithms and database suppliers. Thanks to it data from databases (tables with columns and rows) are transformed into a needed form by the data mining algorithms (cases and their attributes).

DMX queries are used to create structure and model of data mining, in order to learn data models and prediction.

- Creation of a structure data mining in DMX
DMX query for the creation of our mining structure:

```
CreateMiningStructure [CompHTA_Structure]
(
  IDCAS           Long           key,
  anurie          Boolean        Discrete,
  antecedent      Boolean        Discrete,
  angine_de_poitrine Boolean Discrete,
  cat_IMC         Text           Discrete,
  cat_chiffre     Text           Discrete,
  Cat_Age         Text           Discrete,
  convulsion      Boolean        Discrete,
  Syncope         Boolean        Discrete,
  Maux_de_tete   Boolean        Discrete,
  Gene_Respiratoire Boolean Discrete,
```

```

Trouble_oculaire Boolean Discrete,
Vertige Boolean Discrete,
Pollakiurie Boolean Discrete,
Pollyurie Boolean Discrete,
Fatigue Boolean Discrete,
Sexe Text Discrete,
AVC Boolean Discrete,
IRA Boolean Discrete,
Cardio Boolean Discrete
)
    
```

- Addition of a data mining model in a structure

DMX request for the creation of our data mining model is the following:

```

AlterMining [CompHTA_Structure]
ADDMiningModel [MNN_CompHTA]
(
IDCAS key,
anurie ,
antecedent,
angine_de_poitrine,
cat_IMC,
cat_chiffre,
Cat_Age,
convulsion,
Syncope ,
Maux_de_tete ,
Gene_Respiratoire,
Trouble_oculaire,
Vertige,
Pollakiurie,
Pollyurie,
Fatigue,
Sexe,
AVC Predict_only,
IRA Predict_only,
Cardio Predict_only
) using Microsoft_Neural_NetWork
    
```

- Creation of prediction requests

Prediction requests of prediction that we use are prediction requests of singleton type. There is another type of prediction request: request of prediction by the batch used when one wishes to make simultaneously predictions on a set of cases.

Prediction requests which we used are the following:

Request for the prediction of IRA

```

SELECTFLATTENEDPredictHistogram(IRA) as Pavc FROM
MNN_CompHTA "
&" NATURALPREDICTIONJOIN "&" (SELECT ""& maux &"AS
maux_de_tete,"
&Sync &" AS Syncope, ""& trouble &"AS trouble_oculaire,
""
& vertige &"AS vertige, ""& convulsion &"AS convulsion, ""&
angine &"AS angine_de_poitrine, ""& polla &"AS pollakiurie,
""& poly &"AS polyurie, ""& fatigue &"AS fatigue, ""&
antecedent &"AS antecedent, ""& anurie &"AS anurie, ""&
gene &"AS gene_respiratoire, ""& sexe &"AS sexe,
""& catage &"AS cat_Age, ""& catchiffre &"AS cat_chiffre,
""& catimc &"AS cat_IMC.) AS t
    
```

Request for the prediction of cardiac

```

SELECTFLATTENEDPredictHistogram(CARDIO) as Pavc FROM
MNN_CompHTA "
&" NATURALPREDICTIONJOIN "&" (SELECT ""& maux &"AS
maux_de_tete,"
&Sync &" AS Syncope, ""& trouble &"AS trouble_oculaire,
""
& vertige &"AS vertige, ""& convulsion &"AS convulsion, ""&
angine &"AS angine_de_poitrine, ""& polla &"AS pollakiurie,
""& poly &"AS polyurie, ""& fatigue &"AS fatigue, ""&
antecedent &"AS antecedent, ""& anurie &"AS anurie, ""&
gene &"AS gene_respiratoire, ""& sexe &"AS sexe,
""& catage &"AS cat_Age, ""& catchiffre &"AS cat_chiffre,
""& catimc &"AS cat_IMC.) AS t
    
```

Request for the prediction of stroke

```

SELECTFLATTENEDPredictHistogram(AVC) as Pavc FROM
MNN_CompHTA "
&" NATURALPREDICTIONJOIN "&" (SELECT ""& maux &"AS
maux_de_tete,"
&Sync &" AS Syncope, ""& trouble &"AS trouble_oculaire,
""
& vertige &"AS vertige, ""& convulsion &"AS convulsion, ""&
angine &"AS angine_de_poitrine, ""& polla &"AS pollakiurie,
""& poly &"AS polyurie, ""& fatigue &"AS fatigue, ""&
antecedent &"AS antecedent, ""& anurie &"AS anurie, ""&
gene &"AS gene_respiratoire, ""& sexe &"AS sexe,
""& catage &"AS cat_Age, ""& catchiffre &"AS cat_chiffre,
""& catimc &"AS cat_IMC.) AS t
    
```

IV. IMPLEMENTATION

By referring to the modeling of the system, we have designed the application by using Visual Basic 8.0 programming language.

Patient's classes, Case-complication, Factors and Risk gave rise to table of our

database. Allotment has given rise to table of columns.

For a better administration of the system, we inserted a modulus of data management of user's accounts.

In what follows are captures done during the implementation along with extracts of source code program.

• Screenshots

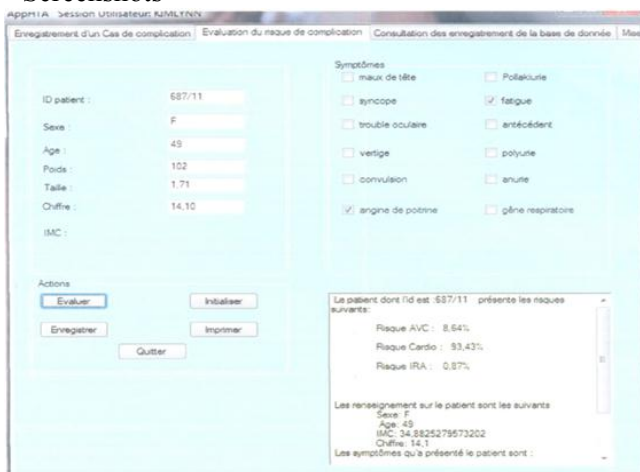


Figure 1: Evaluation of risks of a patient

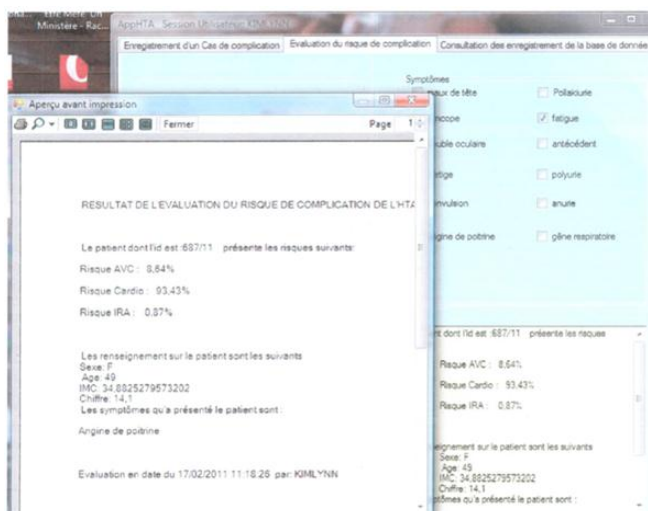


Figure 2: Preview before printing of results of an evaluation of an evaluation

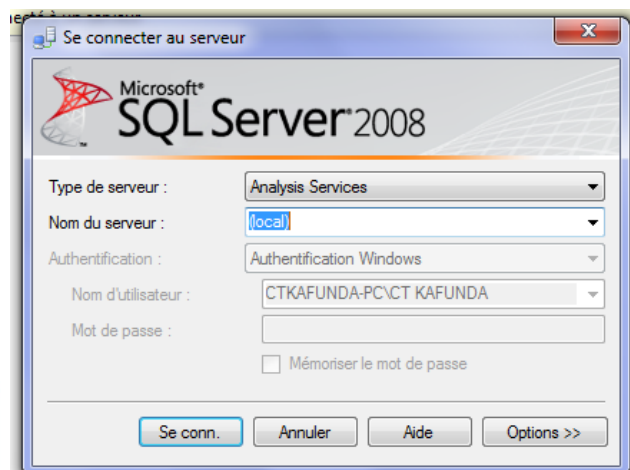


Figure 3: Starting of the server analysis

VI. Extract of sources codes

-Code relatif à la classe Loggin qui permet la génération de la fenêtre de connexion et lancement

```
Public Class Loggin
    Dim code, id As String
    Public use As Integer
    Public acces As String ' variable pour le contrôle des accès à l'application
    Dim i As Integer

    Private Sub UsersBindingNavigatorSaveItem_Click(ByVal sender As System.Object, ByVal e As System.EventArgs)
        Me.Validate()
        Me.UsersBindingSource.EndEdit()
        Me.TableAdapterManager.UpdateAll(Me.UserAppHTADataset)
    End Sub

    Private Sub Loggin_Load(ByVal sender As System.Object, ByVal e As System.EventArgs) Handles MyBase.Load
        'TODO : cette ligne de code charge les données dans la table 'UserAppHTADataset.Users'. Vous pouvez la déplacer ou la supprimer selon vos besoins.
        Me.UsersTableAdapter.Fill(Me.UserAppHTADataset.Users)
        i = 0
    End Sub

    Private Sub Bok_Click(ByVal sender As System.Object, ByVal e As System.EventArgs) Handles Bok.Click
        'on recupere le contenu des textboxes
        If i < 3 Then

            If rbadmin.Checked = True Then
                use = 0
            ElseIf rbfinal.Checked = True Then
                use = 1
            Else
                use = 1
            End If
            'pour la vérification des code d'accès entre les user et l'administrateur
            'normalement on doit lire dans un fichier où se trouve les code d'accès et les utilisateurs
            'il ya deux fichiers différents users et admin

            Dim rep As Integer
            id = txtid.Text
            If (id = "") Then
                MsgBox("vous n'avez rin entré comme identifiant veuillez le faire sinon l'opération sera annuler")
                i = i + 1
            Exit Sub
            End If
        End Sub
    End Class
```

```
code = txtcode.Text
If (code = "") Then
MsgBox("vous n'avez rin entré comme code, veuillez le faire sinon
l'opération sera annuler")
i = i + 1
Exit Sub
End If
If use = 0 Then
'on verifi que code et id ce trouve dans la table des administrateur
carseul eux on accès à certain module

rep = Me.UsersTableAdapter.recherche(id, code, "ADMIN") 'rechercher
est une table qui renvoi le nombre d'enregistrement correspondant
If rep = 0 Then
acces = "no"
ElseIf rep = 1 Then
acces = "ok"
End If
ElseIf use = 1 Then
'ici on cherche parmi les simple utilisateur
rep = Me.UsersTableAdapter.recherche(id, code, "FINAL") 'rechercher
est une table qui renvoi le nombre d'enregistrement correspondant
If rep = 0 Then

acces = "no"
ElseIf rep = 1 Then
acces = "ok"
End If
End If
'on verifie la valeur de acces pour lancer ou pas les action
If (acces = "ok") Then
Me.UsersTableAdapter.majdate(Date.Now, id) 'on enregistre le dernier
accès
Form1.Show()
Me.Hide()
Else
MsgBox("Vous n'avez pas les droits requis pour accéder à ce module")
End If
End If
End Sub
Private Sub rbadmin_CheckedChanged(ByVal sender As System.Object,
ByVal e As System.EventArgs) Handles rbadmin.CheckedChanged

End Sub
Private Sub BAnn_Click(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles BAnn.Click
'on verrifie que l'utilisateur vuet reellment quitter l'application
Dim rep
rep = MsgBox("Etés vous sûr de vouloir quitter l'aplication",
MsgBoxStyle.OkCancel)
If MsgBoxResult.Ok = rep Then
'on quitte l'aplication
Application.Exit()
End If
End Sub
End Class
```

-Code relatif à la classe principale de l'application

*Importation des bibliothèques nécessaires

```
Imports Microsoft.AnalysisServices.AdomdClient
Imports Microsoft.AnalysisServices.AdomdClient.MiningStructure
Imports Microsoft.AnalysisServices.AdomdClient.MiningModel
Imports Microsoft.AnalysisServices.System.DataMining
Imports System.IO
Imports System.Drawing.Printing
Imports System.Diagnostics
```

Public Class Form1

```
'déclarationdes variable
'variable pour l'impression
Dim sp As StreamReader
Dim pf As Font
Dim entête As String
'variable pour le controle d'accès
Dim acces As String
'variable correspondants au checkbox
Dim mau, sync, trouble, vertige, convulsion, angine, polla, poly,
fatigue, antecedent, anurie, gene As Boolean
'fin variable checkbox
'variable chaîne de caractère pour récupérer les symptômes du
patient
Dim symptome As String
```

'debut variable cas complication

```
Dim avc, cardio, ira As Boolean
Dim act As Char()
Dim cas, idmof As Integer
'fin variable
```

'variable pour enregistrement du cas et des patients

```
Dim idpatient, nomp, pncmp As String
Dim idcas As Integer
Dim sexe As Char
Dim taille As Double
Dim poids As Double
Dim age As Integer
Dim imc As Double
Dim chiffre As Double
Dim datecas, pdate As Date 'pdate est la date de naissance du
patient
```

'varibale pour la partie visualisation et modification

```
Dim ajout, modif, supp As Boolean
Dim tpat, tcas, trisq, tfact As Boolean
```

'fin variable en registrement cas

```
Dim i, j, k As Integer 'ceux sont les competur d'itération dontnous
allons nous servir
'variable pour l'evaluation des risques
Dim Pavc, Pira, Pcardio As Double 'variable qui von garder la proba
```

```
Dim catage, catimc, catchiffre As String 'variable qui contiendra les
catégorie pour les variable numérique
Dim aide As Boolean 'une variable d'aide pour le test ou
verification
'if faudra utiliser le Dataset2 pour l'évaluation

' les donnée pour l'administration et la mise à jour des modesl de
donnée

'variable chaîne de caractère qui contient le text de la requête il
y en aura trois
Dim reqa, reqi, reqc As String
'une sorte de data reader pour les bases de donnée analys

Private Sub CompHTABindingNavigatorSaveItem_Click(ByVal sender As
System.Object, ByVal e As System.EventArgs)
Me.Validate()
Me.CompHTABindingSource.EndEdit()
Me.TableAdapterManager.UpdateAll(Me.SourceHTADataset)
End Sub

Private Sub Form1_Load(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles MyBase.Load
'TODO : cette ligne de code charge les données dans la table
'UserAppHTADataset.Users'. Vous pouvez la déplacer ou la supprimer
selon vos besoins.
Me.UsersTableAdapter.Fill(Me.UserAppHTADataset.Users)
'TODO : cette ligne de code charge les données dans la table
'SourceHTADataset2.Risque'. Vous pouvez la déplacer ou la supprimer
selon vos besoins.
Me.RisqueTableAdapter.Fill(Me.SourceHTADataset2.Risque)
'TODO : cette ligne de code charge les données dans la table
'SourceHTADataset2.Facteurs'. Vous pouvez la déplacer ou la
supprimer selon vos besoins.
Me.FacteursTableAdapter.Fill(Me.SourceHTADataset2.Facteurs)
'TODO : cette ligne de code charge les données dans la table
'SourceHTADataset2.Patients'. Vous pouvez la déplacer ou la
supprimer selon vos besoins.
Me.PatientsTableAdapter1.Fill(Me.SourceHTADataset2.Patients)
'TODO : cette ligne de code charge les données dans la table
'SourceHTADataset2.CompHTA'. Vous pouvez la déplacer ou la supprimer
selon vos besoins.
Me.CompHTATableAdapter2.Fill(Me.SourceHTADataset2.CompHTA)
'TODO : cette ligne de code charge les données dans la table
'SourceHTADataset1.Patients'. Vous pouvez la déplacer ou la
supprimer selon vos besoins.
Me.PatientsTableAdapter.Fill(Me.SourceHTADataset1.Patients)
'TODO : cette ligne de code charge les données dans la table
'SourceHTADataset1.CompHTA'. Vous pouvez la déplacer ou la supprimer
selon vos besoins.
Me.CompHTATableAdapter1.Fill(Me.SourceHTADataset1.CompHTA)
'TODO : cette ligne de code charge les données dans la table
'SourceHTADataset.CompHTA'. Vous pouvez la déplacer ou la supprimer
selon vos besoins.
Me.CompHTATableAdapter.Fill(Me.SourceHTADataset.CompHTA)
```

54

CONCLUSION

This work on data mining allowed to develop an information system of assessment of the risk of complication of a high blood pressure at patients. This result contributes to the development of data contained in information systems in the hospital complexes and in general in businesses, a goal reached by datamining.

The use of data mining methods in medicine is growing rapidly. As in other areas, the availability of large databases needs the use of this technique. Most of the published papers focus on risk factors for home accidents, diabetes, suicide, hospital infections, etc. This work also brings a contribution to this area.

BIBLIOGRAPHY

- [1] **AUPETIT A**: *les réseaux de neurones artificiels ; Laboratoire d'Algorithmique 2004.*
- [2] **BISHOP C.M**: *Neuronal Network for Pathernrecognition, Clarendon Press, oxford, 1995.*
- [3] **CHAMBO J**: *Concepts d'informatique décisionnelle ; cours université de Provence, 2010.*
- [4] **LEMAIRE V**: *les réseaux de neurones multicouches, cours université de Reune, 2007.*
- [5] **KASENDE TSHITUTA I.M** : *Applications du Data Mining, Mémoire de Licence en Sciences Informatiques, Kinshasa, 2010.*
- [6] **MACLENNAN J, TANG, CRIVAT B,** :*Data mining with Microsoft SQL Server 2008, Willey 2009.*
- [7] **MBUYI MUKENDI E** : *Intelligence artificielle approfondie, Cours Université de Kinshasa, 2010.*
- [8] **MBUYI MUKENDI E** : *Data Minig and Network neuronal, Extracting knowledge from high pressure data patients, IJCSI, Volume 9, Issue 3 , 2012*
- [9] **PEREZ-URIBE A** : *Réseau de neurones artificiels et apprentissage non-supervisé, Hieg-vd 2009.*
- [10] **RAKOTOMALALA R** : *Introduction au datamining et arbre de décision, Revue Modulad n°33, 2005.*
- [11] **RAKOTOMALALA R** : *Introduction au datamining cours Rouen 2007.*
- [12] **SAPORTA G** : *Datamining : une nouvelle façon de faire des statistiques, CNAM, 2002.*
- [13] **SCHNEDER R. GIBSOND** :*Microsoft SQL Server 2008 for Domnies ; Willey 2008.*
- [14] **SERNA ENICAS M** : *Entrepôt de données pour l'aide à la décision médicale : conception et expérimentation, Thèse de Doctorat, Université Joseph Fourier, 2005.*
- [15] **TRUFFEYS** : *Data mining et Scornig ; édition Dunod, 2002.*
- [16] **TRUFFEYS** : *Data mining et statistique décisionnelle, édition Technip 2007.*