Big data decision making: an application activity

Carolyn LaMacchia, Ph.D.
Bloomsburg University of Pennsylvania

John Rude, Ph.D., CPA
Bloomsburg University of Pennsylvania

Elaine Dolak, M.Ed.
Bloomsburg University of Pennsylvania

ABSTRACT

This research introduces a big data approach to decision making in higher education through an activity that applies this new paradigm along with the familiar decision table process to a loan application approval problem. Much of business decision making has been based on the mechanization of the decision table process. Computer systems offer a solution by comparing situational data with rules developed by subject matter experts. Organizations are adopting a new paradigm in decision making and support, based on Big Data, that leverages advances in data storage and computing power. This alternative decision making process is predictive analytics, which recommends a decision based on the examination of historical and current situational data. Business professionals should recognize and comprehend the logic that drives decision support systems encountered during their careers. This research addresses the introduction of this paradigm in decision making in higher education through a hands-on application activity. Decision tables and predictive analytic processing are discussed. An application activity is presented that incorporates decision table and predictive analytic decision support techniques for a loan application approval process. This activity is an opportunity to develop an awareness of decision making techniques in undergraduate and graduate students.

Keywords: Information Technology, Business Education, Decision Tables, Predictive Analytics, Decision Support

INTRODUCTION

A number of decision support systems used to solve both simple and complex problems are based on models embodying the knowledge of domain experts (Brown and Chandrasekaran, 2014). Systems can harness data and computation power to reliably provide quantitative results that support the power of human intelligence. Predictive analytics, a category of decision-support systems, uses statistical and computational methods along with past and present data to form models regarding what may happen in the future. Methods employed in these systems are based in the theory of classification trees, regression analysis, machine learning, and neural networks (Hoffer, Ramesh, and Topi, 2016). Predictive analytics can be used to improve an organization's understanding to make better business decisions (Parr-Rud, 2012).

The use of predictive analysis in business organization is dramatically increasing. (DeAngelis, 2015). Organizations are leveraging data and predictive analytic tools to provide insights to support corporate strategy (DeAngelis, 2015). The National Association of Colleges and Employers conducted a survey of employers and report the need e to hire new college graduates that can make decisions and solve problems (Adams, 2014). In addition, employers prefer new hires with the technical knowledge that support modern decision making techniques (Adams, 2014). This paper addresses the introduction of predictive analytics in higher education to prepare undergraduate and graduate business majors for decision making techniques and technology tools that they may encounter in their careers. Decision table processing and predictive analytics processing are discussed. There is an undeniable attraction in the rigor that computational modelling can bring to the description of learning (Laurillard, 2013). Acknowledging the impact of hands-on learning, an assignment is presented that uses decision table and predictive analytics to support the decision making process for loan approval, a familiar business decision.

Decision Tables

Decision tables are a modelling tool implemented to illustrate and understand complex procedures and policies. Decision tables are useful for the presentation of logic in a clear tabular form as opposed to a narrative form (Gelenbe, 2014). General Electric, the Sutherland Corporation, and the United States Air Force attempted to model a complex file maintenance project using flowcharts and traditional narratives (LaFleur, 1971). Six labor-years of effort was placed in the project but they failed to define the problem (LaFleur, 1971). Yet, four analysts using decision tables successfully defined the problem in less than four weeks (LaFleur, 1971). Business analysts often work with a subject matter expert to examine the business problem/opportunity and develop a model within a decision table. Once complete, the decision table is typically passed to programmers for implementation in a computer system.

A decision table contains all possible conditions and the appropriate action for each condition. Figure 1 (Appendix) illustrates the typical decision table. Conditions to evaluate are located in the upper half and possible actions to be taken in the lower half. Each column in the table represents a rule. All possible combinations of values in the columns are entered in the table.

Predictive Analytics

With processing logic based on decision table modeling, the rules and parameters are defined for a computer. Predictive analytic tools are a new paradigm where the computer uses (lots of) data to define the rules and parameters. The model for this tool has its roots as far back as the late 1950's, when IBM's Arthur Samuel used a new approach to play checkers with the computer (McCarthy, 1990). After initially programming the computer with the fundamental checker moves, Arthur Samuel discovered that the computer also needed to "learn" strategy. Additional data was collected during a series of many games: each opponent's move and the opponent who won the game. This, in effect, "taught" strategy to the computer. Using this new technique, when presented with a situation, the computer reviews its database of possible moves and chooses the one that has the highest probability of leading to a win (McCarthy, 1990).

The methods for predictive analytics systems are not new. The models are based on well-established statistical and computation methods. What has changed is the ease which the processes are applied to practical organization questions including sales, procurement, production system controls, warranties, and customer relationship management. Organizations have recognized the capabilities of the predictive analytics approach. It is important to note that while the models can achieve a high predictive accuracy rate, the reasoning behind how the decision is reached is not readily available. Licklider, 1960 seminal research describes how the symbiosis of the human and computer, where human knowledge works with computing power, has the potential to create the best decision making scenarios (Licklider, J. C. R., 1960)

DECISION SUPPORT ACTIVITY

The purpose of this activity is to explore the application of two different decision support techniques for the same business problem. The business activity is a familiar business decision - the loan approval process. This hands-on exercise provides undergraduate and graduate business with the experience of applying decision table decision support and predictive analysis decision support to a business problem. The activity is presented in two parts. Part 1 presents the application of the decision table method and Part 2 presents the application of the predictive analytics method.

Part 1: Decision Support Activity based on Decision Table

The details of the decision support activity based on the decision table method is presented in this section of the paper. The decision support activity is presented here. The suggested solution appears in the Appendix. The loan application processing policy uses a rule-based module to reach a decision. The rule-based module evaluates loan application data and decides whether to Approve (A), Reject (R), or Send-to-Supervisor for future evaluation (S).

Loan Specialists (subject-matter experts) have determined that data should be derived for each loan application. Four factors are generated:

1. Credit Rating 1 (from National Credit Bureau) with values Unsatisfactory (U) and Satisfactory (S). The National Credit Bureau, a fictional organization, provides credit scores on loan applicants. Scores at or above 650 are considered satisfactory; score below 650 are considered unsatisfactory.

2. Credit Rating2 (from Trans-Continental Assoc.) with values Unsatisfactory (U) and Satisfactory (S). Trans-Continental Associate, a fictional organization, provides credit scores on loan applicants. Scores at or above 630 are considered satisfactory; score below 630 are considered unsatisfactory.
3. Income-to-loan-amount-ratio with values High (H), Medium (M), and Low (L).
A high income-to-loan-amount ratio demonstrates an acceptable balance between income and the borrowed funds. Conversely, a low ratio can signal that an individual is requesting too high of a loan amount when considered in the context of the amount of their total income.
4. Debt-to-income-ratio with values High (H), Medium (M), and Low (L).
A low debt-to-income ratio demonstrates an acceptable balance between debt and income. Conversely, a high ratio can signal that an individual has too much debt when considered in the context of the amount of their total income.

Loan Specialist have agreed upon evaluation criterion the four factors. The business rules are as follows:

1. An application is "Rejected" if either of the two Credit Ratings (from National Credit Bureau or Trans Continental Assoc.) is "Unsatisfactory".
2. An application is also "Rejected" if either the Income-to-loan-amount-ratio is "Low" or the Debt-to-income-ratio is "High".
3. An application with both Credit Ratings as "Satisfactory" is "Approved" if the Income-to-loan-amount-ratio is "High" and the Debt-to-income-ratio is "Low".
4. All other cases are evaluated on an individual basis by a Loan Specialist.

Part 1 Deliverables

Step 1 may be completed by business students in undergraduate or graduate technology course. Steps 2 and 3 are optional and may be used in introductory information technology courses.
Step 1: Represent the logic of the above loan application processing policy by constructing a decision table. Simplify the decision table as much as possible using "don't-care" states. Ensure that the table is correct and complete. This is the portion of the process that is typically developed by business analysis working with a subject matter expert. A suggested solution may be found in Part 1, Step 1 - Suggested Solution (Appendix).
Step 2: Specify the logic as IF..THEN Rules. Avoid using ELSE Statements. This portion of the process is usually generated for the programmer to implement the decision table logic. A suggested solution may be found in Part 1, Step 2 - Suggested Solution (Appendix).
Step 3: Create a program that accepts the four decision variables and determines the loan approval status. The programmer creates the mechanized solution based on the If..Then rules. A suggested solution may be found in Part 1, Step 3 - Suggested Solution (Appendix).

Part 2: Decision Support Activity based on Predictive Analytics

The financing company has been collecting data on loan recipients for a number of years. Data includes a number a variables that describe the life style and habits of the individual where

a loan was approved. Most importantly, the data reveals whether the loan customer paid back or defaulted on the loan. This loan application processing policy analyzes current data in the context of historical data to reach a decision.

Part 2 Deliverables

1. Download and install JustNN, a free, simple to use neural network which introduces predictive analytic models. Recommendation: install on all computers in lab. Download is available from: www.justnn.com

2. Complete the tutorial exercises found in Getting Started these include: XoR function, Color Circle, and Races. Students are introduced to the tool operation and concept by completing the tutorial exercises.
3. Describe the purpose of predictive analytics processing. Describe training. (optional)
4. Evaluate a borrower potential for repaying a loan based on predictive analytics decision making method. The directions are as follows:
     a. Locate a credit screening data file. Prepare file and store in spreadsheet format. Add descriptive column headers (field name) to each field. Recommend variables that you hypothesize may predict credit worthiness.
     Recommendation:
     U.C. Irvine Machine Learning Repository includes a Credit Screening file that is available for free download (U.C. Irvine Machine Learning Repository). To prepare a file for predictive analytics processing, students should access the following three files:
          1. credit.names – contains descriptive information about the file.
          2. crx.data – contains the data.
          3. crx.names – contains detailed descriptive of the data including field name.
     b. Analyze data using JustNN by creating and training the network with a same file.
     Recommendation:
     If the Credit Screening file is used for this exercise, consider using only the first 150 records for fast performance.
     c. Describe advantages of a predictive analytics approach.
     d. Describe weaknesses of a predictive analytics approach.

CONCLUSION

     Thriving in a highly competitive business environment depends on quality business decisions. Organizations support decisions with mechanized applications based on both decision table processing and predictive analytics processing. Business professionals should recognize and comprehend the logic that drives decision support systems encountered during their careers. This research addresses the introduction of this paradigm in higher education. A hands-on activity has been developed that recognizes that learning should be viewed as a series of activities that develop skills and capabilities. Learning should not just be about acquiring high-level knowledge; it is about how students handle that knowledge (Laurillard, 2013). Decision tables and predictive analytic processing are demonstrated through an activity that illustrates a common business decision. This exercise is an opportunity to develop an awareness of decision making techniques in undergraduate and graduate students.

APPENDIX

Figure 1 - Typical Decision Table

Conditions (decision variables)	Rules for each possible value of the variables
Decision (actions)	Decision is made by using the rule to evaluate the value of condition variable

Part 1, Step 1 - Suggested Solution

Rules ------------------------>	1	2	3	4	5	6	7	8
Decision Variables								
Credit Rating 1 (U/S)	U	~	~	~	S	S	S	S
Credit Rating 2 (U/S)	~	U	~	~	S	S	S	S
Income/Loan Amount Ratio (H/M/L)	~	~	L	~	H	H	M	M
Debt/Income Ratio (H/M/L)	~	~	~	H	L	M	L	M
Decision: Loan Approval Process (A/R/S)	R	R	R	R	A	S	S	S

Legend

U	Unsatifactory
S	Satisfactory
L	Low
M	Medium
H	High
~	Don't Care
R	Reject
A	Approve
SS	Send to Supervisor

Part 1, Step 2 - Suggested Solution

Rule										
1	If Credit Rating 1 (CR1)			is	U	then		Loan Approval Process	is	R
2	If Credit Rating 2 (CR2)			is	U	then		Loan Approval Process	is	R
3	If Income/Loan Amount Ratio (I/L)			is	L	then		Loan Approval Process	is	R
4	If Debt/Income Ratio (D/I)			is	H	then		Loan Approval Process	is	R
5	If CR1 is "S" and CR2 is "S" and I/L Ratio is "H" and D/I Ratio is "L"					then		Loan Approval Process	is	A
6	If CR1 is "S" and CR2 is "S" and I/L Ratio is "H" and D/I Ratio is "M"					then		Loan Approval Process	is	SS
7	If CR1 is "S" and CR2 is "S" and I/L Ratio is "M" and D/I Ratio is "L"					then		Loan Approval Process	is	SS
8	If CR1 is "S" and CR2 is "S" and I/L Ratio is "M" and D/I Ratio is "M"					then		Loan Approval Process	is	SS

Part 1, Step 3 - Suggested Solution

The suggested solution uses Microsoft Excel® and macro processing. The Input Value cells include data validation to verify the values at the point of data entry.

Loan Approval Process		Directions: Input Decision Variables, Click "Determine Loan Approval"

Decision Variables	Input Value
Credit Rating 1 (U/S)	
Credit Rating 2 (U/S)	
Income/Loan Amount Ratio (H/M/L)	
Debt/Income Ratio (H/M/L)	
Decision: Loan Approval Process (A/R/S)	

Legend	
Credit Ratings	
U	Unsatisfactory
S	Satisfactory
Ratios	
H	High
M	Medium
L	Low
Decision	
A	Accept
R	Reject
S	Send to Supervisor

Determine Loan Approval

Clear Decision Variables

Macro code:

```
Sub DetermineLoanApproval()
' DetermineLoanApproval macro triggered by button
Dim num1 As Integer
num1 = 0
'Rule 1: Credit Rating 1 is Unsatisfactory.
'*** Reject application
If Range("B5") = Range("F5") Then
    Range("B9").Value = Range("F13")
    num1 = 1
End If
'Rule 2: Credit Rating 2 is Unsatisfactory.
'*** Reject application
If Range("B6") = Range("F5") Then
    Range("B9").Value = Range("F13")
    num1 = 1
End If
'Rule 3: Income/Loan Ratio is Low.
'*** Reject application
If Range("B7") = Range("F10") Then
    Range("B9").Value = Range("F13")
    num1 = 1
End If
'Rule 4: Debt/Income Ratio is High.
'*** Reject application
```

```
If Range("B8") = Range("F8") Then
   Range("B9").Value = Range("F13")
   num1 = 1
End If
'Verify Input'
'Rule 0: Is Data Entered?"
If num1 = 0 Then
   If Range("B5") = "" Or Range("B6") = "" Or Range("B7") = "" And Range("B8") = "" Then
   Range("B9").Value = "?"
   num1 = 1
   End If
End If
'Rule 5: Credit Rating 1 and Credit Rating 2 are Satisfactory,
'Income/Loan is Ratio is High,
'Debt/Income is Low.
'*** Accept Application
If num1 = 0 Then
   If Range("B5") = Range("F6") And Range("B6") = Range("F6") And Range("B7") =
Range("F8") And Range("B8") = Range("F10") Then
   Range("B9").Value = Range("F12")
   num1 = 1
   End If
End If
'Rule 6: Credit Rating 1 and Credit Rating 2 are Satisfactory,
'Income/Loan is Ratio is High,
'Debt/Income is Medium.
'*** Send Application to Supervisor
If num1 = 0 Then
   If Range("B5") = Range("F6") And Range("B6") = Range("F6") And Range("B7") =
Range("F8") And Range("B8") = Range("F9") Then
   Range("B9").Value = Range("F14")
   num1 = 1
   End If
End If
'Rule 7: Credit Rating 1 and Credit Rating 2 are Satisfactory,
'Income/Loan is Ratio is Medium,
'Debt/Income is Low.
'*** Send Application to Supervisor
If num1 = 0 Then
   If Range("B5") = Range("F6") And Range("B6") = Range("F6") And Range("B7") =
Range("F9") And Range("B8") = Range("F10") Then
   Range("B9").Value = Range("F14")
   num1 = 1
   End If
End If
'Rule 8: Credit Rating 1 and Credit Rating 2 are Satisfactory,
```

```
'Income/Loan is Ratio is Medium,
'Debt/Income is Medium.
'*** Send Application to Supervisor
If num1 = 0 Then
    If Range("B5") = Range("F6") And Range("B6") = Range("F6") And Range("B7") =
Range("F9") And Range("B8") = Range("F9") Then
    Range("B9").Value = Range("F14")
    numb1 = 1
    End If
End If
If num1 = 0 Then
    Range("B9").Value = "?"
    num1 = 1
End If
End Sub
Sub ClearDecisionVariables()
' Clear Decision Variables macro triggered by button.
    Range("B5").Select
    ActiveCell.FormulaR1C1 = ""
    Range("B6").Select
    ActiveCell.FormulaR1C1 = ""
    Range("B7").Select
    ActiveCell.FormulaR1C1 = ""
    Range("B8").Select
    ActiveCell.FormulaR1C1 = ""
    Range("B9").Select
    ActiveCell.FormulaR1C1 = ""
    Range("B5").Select
End Sub
```
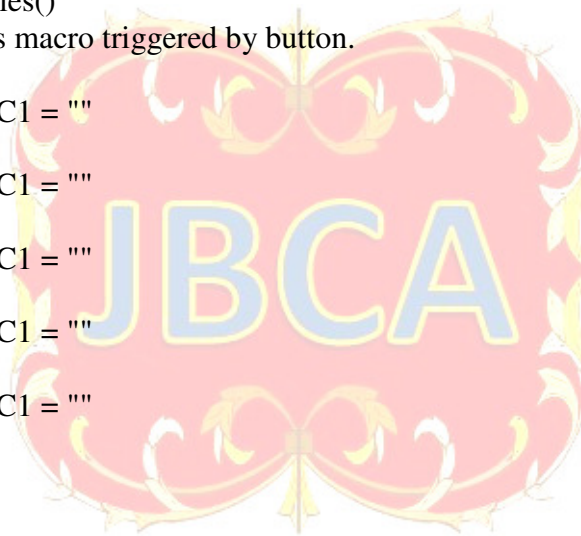
References

Adams, S. (2014). The 10 Skills Employers Most Want in 2015 Graduates. Forbes. Retrieved August 1, 2015 from http://www.forbes.com/sites/susanadams/2014/11/12/the-10-skills-employers-most-want-in-2015-graduates/#55d133af19f6

Brown, D. C., & Chandrasekaran, B. (2014). *Design problem solving: knowledge structures and control strategies*. Morgan Kaufmann.

DeAngelis, S. F. (2015). Predictive Analytics becoming a Mainstream Business tool. *Enterra Solutions*. Retrieved August 1, 2016 from http://www.enterrasolutions.com/2015/04/predictive-analytics-becoming-a-mainstream-business-tool.html

Gelenbe, E. (2014). *Analysis and Synthesis of Computer Systems.* John Wiley & Sons, Hoboken, NJ, USA.

Hoffer, J.A., Ramesh, R., & Topi, H. (2016). *Modern Database Management*, 12e. Pearson.

LaFleur, T.G (1971). Decision tables: a tool for documenting logical condition relationships. *ACM SIGPLAN Notices – Special issue on decision table*. 6(18): 9-12.

Laurillard, D. (2013). *Rethinking University Thinking: A Conversational Framework for the Effective Use of Learning Technologies*. Routledge.

Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE transactions on human factors in electronics*, (1): 4-11.

Parr-Rud, O. (2012). Drive Your Business With Predictive Analytics. *SAS Institute*, available at www.sas.com/en_us/whitepapers/dirve-your-business-with-predictive-analytics-105620.html

McCarthy, J. (1990). In Memoriam - Arthur Samuel: Pioneer in Machine Learning. *AI Magazine*, 11(3): 10 – 11.

U.C. Irvine Machine Learning Repository. (n.d.) Credit Approval. Retrieved August 11, 2016 from http://archive.ics.uci.edu/ml/datasets/Credit+Approval

Watson, H. (2014). Tutorial: Big Data Analytics: Concepts, Technologies, and Applications. *Communication of the AIS,* 34,65: 1247-1268.