Neural Networks and Reducts Applied to Credit Card Fraud Detection

Jordan Moreno-Lacalle

Mentor: Dr. Wellington

I. Problem Description

Does the use of a neural network in conjunction with feature selection by finding reducts result in satisfactory performance for credit card fraud detection?

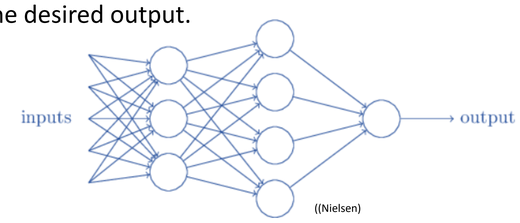
II. Literature Review

ANN: An ***artificial neural network*** (or simply, *neural network*) is "a set of interconnected nodes designed to represent [the] functioning of [the] human brain" (Chaudhary).

Each neuron in a neural network takes some *n* number of inputs *X.* Each input X_i has an associated weight W_i. Each neuron also has a bias (a measure of the neuron's affinity for activation).

The output or activation of a neuron is determined by passing the weighted sum of its inputs plus the bias to the activation function. A popular activation function is the sigmoid activation function which restricts the output of a neuron to the range of real numbers from 0 to 1.

Neural networks are trained by minimizing the error measured by a cost function. The cost function measures how the neural network's output compares to the desired output.



((Nielsen)

A common method for minimizing this error is gradient descent, which is implemented by the ***backpropagation*** algorithm.

Backpropagation iteratively improves a neural network model by computing the error for the current set of weights and biases and backpropagating this error through each layer. This adjusts the weights and biases to approach a more predictive model.

A ***reduct*** is a minimal set of attributes that preserves the ability to properly classify objects. Reducts originate from rough sets theory.

Finding reducts requires discrete attributes, so continuous attributes will need to be discretized. For the purposes of this project, simple binning would not be very effective.

Some more effective methods of discretization are based on the ***Minimum Description Length (MDL) Principle.*** Methods based on this principle exploit the regularities or patterns in data to describe the data using fewer bits. A reputable MDL-based method is Fayyad and Irani's method detailed in "Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning" which uses an MDL-based criterion to determine whether or not discretization should continue.

III. Primary Objective

To evaluate the performance of a neural network for credit card fraud detection that uses reducts (from Rough Sets Theory) for feature selection, as proposed in "Credit Card Fraud Detection Using Rough Sets and Artificial Neural Network" (Gupta, et. al.). (1 Person-Month)

Governing propositions:

- The neural network will be trained and evaluated using the data provided in the Kaggle dataset "Credit Card Fraud Detection" provided by the Machine Learning Group of the Free University of Brussels. https://www.kaggle.com/mlg-ulb/creditcardfraud
- Performance will be evaluated primarily by confusion matrices.

		Predicted Class: Fraudulent	Predicted Class: Legitimate
Actual Class	Fraudulent	TP	FN
	Legitimate	FP	TN

Example of a confusion matrix for this project. True positive, false positive, true negative, and false positive values are represented by TP, FP, TN, and FN respectively.

Selectivity: The proportion of transactions that were predicted as legitimate and are actually legitimate.
→ TN/ (FP+TN)

Positive Predictive Value: The proportion of transactions that were predicted as fraudulent and are actually fraudulent.
→ TP/ (TP+FP)

IV. Solution Description

- Tools: JAVA, WEKA
- Kaggle Credit Card Fraud dataset
  - 284,807 transactions
    - 492 are fraudulent → 0.17% fraud ratio
- Select features by finding reducts (will need to discretize data for this step) and preprocess the data as otherwise necessary.
- Develop a program that trains the network on the dataset using several reducts and record data about the classifications. Change the number of hidden layers to more effectively evaluate performance.
- Use 10-way cross validation throughout the trials (train 90% test 10%)

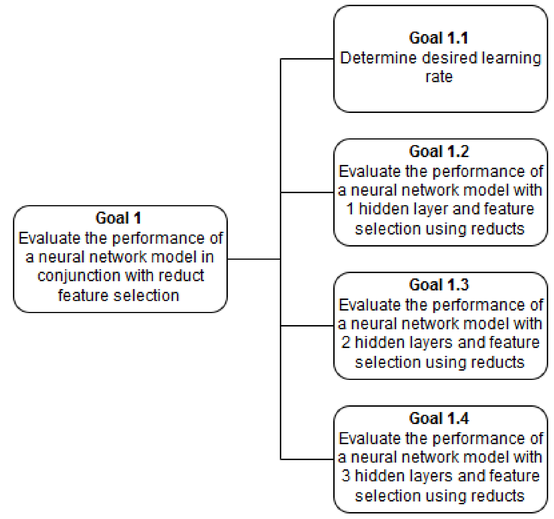
V. Hypotheses

H1: Using reducts for feature selection in conjunction with a neural network will result in satisfactory accuracy (>= 95%) for credit card fraud detection.

H2: Using reducts for feature selection in conjunction with a neural network will result in satisfactory selectivity (>= 90%) for credit card fraud detection.

H3: Using reducts for feature selection in conjunction with a neural network will result in a satisfactory positive predictive value (>= 90%) for credit card fraud detection.

VI. Goal Tree



VI. Experiment Design

Block Design				
10-Way Cross Validation		Number of Hidden Layers		
		1	2	3
Reduct	R1	x	x	x
	R2	x	x	x
	R3	x	x	x

Chaudhary, Khyati, Jyoti Yadav, and Bhawna Mallick. "A Review of Fraud Detection Techniques: Credit Card." *International Journal of Computer Applications* 45.1 (2012): 975–8887. Web. 18 Feb. 2018.

Fayyad, Usama, and Keki Irani. "Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning." *In Proceedings of the 13th International Join Conference on Artificial Intelligence. 1993.*

Grunwald, Peter. "A Tutorial Introduction to the Minimum Description Length Principle." *Advances in Minimum Description Length: Theory and Applications.* MIT Press, 2005. Web. 28 Apr. 2018.

Komorowski, Jan, Lech Polkowski, and Andrzej Skowron. "Rough Sets: A Tutorial." PDF. Web. 12 Apr. 2018. http://eecs.ceas.uc.edu/~mazlack/dbm.w2011/Komorowski.RoughSets.tutor.pdf

Nielsen, Michael A. "Chapter 1 - 3." *Neural Networks and Deep Learning*. Determination Press, 02 Dec. 2017. Web. 15 Feb. 2018.