NEURODECISION(TM): A MATHEMATICALLY RIGOROUS METHOD FOR DEVELOPING NEURAL NETWORKS THAT RESOLVES THE ADVERSE ACTION CODE ASSIGNMENT PROBLEM IN RISK MODELING

Michael McBurnett
Equifax Data and Analytics
April 2018

Executive Summary

- Research has shown that Neural Networks demonstrate potential performance gains over traditional scorecards.
- Regulatory hurdles have precluded use for actionable decisioning.
- We have developed a new Neural Network Model Development Process that facilitates logical reason code generation.
- The process has been approved internally by Equifax legal and compliance to satisfy relevant regulations.
- NeuroDecision™ is patent pending.

Agenda

- Existing scoring regulations
- The case for Neural Networks
- Regulatory challenges with unconstrained Neural Networks
- NeuroDecision Technology: A Neural Network framework satisfying regulatory and business requirements

EXISTING SCORING REGULATIONS

All Credit Scoring Systems Must Conform to FCRA & Reg B

- OCC, CFPB, Dodd-Frank, and FCRA all point to the ECOA Regulation B “notice when adverse action is taken” requirement and FCRA Section 609(f) “Disclosure of Credit Scores”.
- A credit scoring system must be an “empirically derived, demonstrably and statistically sound, credit system.”[1]
- It must meet the requirements written in the Equal Credit Opportunity Act (Regulation B), Title 12, Chapter II, Subchapter A, Part 202(9)(b)(2):
  - It provides a number of specific reasons (4) associated with adverse action
  - The description of the reason is extracted from our current list of adverse action codes
  - It is based only on those attributes considered by the credit scoring system
  - The method for selecting adverse action codes is substantially similar to the methods described in the written regulations
- FCRA Section 609(f)(2)(A): The term “credit score” means a numerical value or a categorization derived from a statistical tool or modeling system …;
- FCRA Section 609(f)(1)(C): all of the key factors that adversely affected the credit score of the consumer in the model used, the total number of which shall not exceed 4, subject to paragraph (9);
- FCRA Section 609(f)(2)(B): The term "key factors" means all relevant elements or reasons adversely affecting the credit score for the particular individual, listed in the order of their importance based on their effect on the credit score.

Logistic Regression is an Example of a Scoring Method that Conforms to FCRA & Reg B

- Logistic Regression Adverse Action Codes are generated from a linear model
  - The model is empirically derived from a population of creditworthy and non-creditworthy applicants
  - The model is developed for the purpose of evaluating the creditworthiness of applicants with respect to the legitimate business purposes of the creditor
  - The model is developed using accepted statistical principles and methods
  - The model is periodically reevaluated and adjusted as necessary
- In a Logistic Regression Model, the impact of a change in any variable is the same for every consumer
- Therefore, Key Factors affecting the model are easily identifiable

Logistic Regression Scorecard Development

- Logistic regression models the probability $p = P(Y = 1)$ of a binary random variable Y:

$$\log\left(\frac{p}{1-p}\right) = f(X_1, \ldots, X_n) = X\beta = \beta_0 + X_1\beta_1 + \cdots + X_n\beta_n$$

- so that

$$p = \frac{1}{1+\exp(-X\beta)}.$$

- Because of the additive nature of logistic regression on the log-odds scale, the maximum points lost only depends upon X_i:

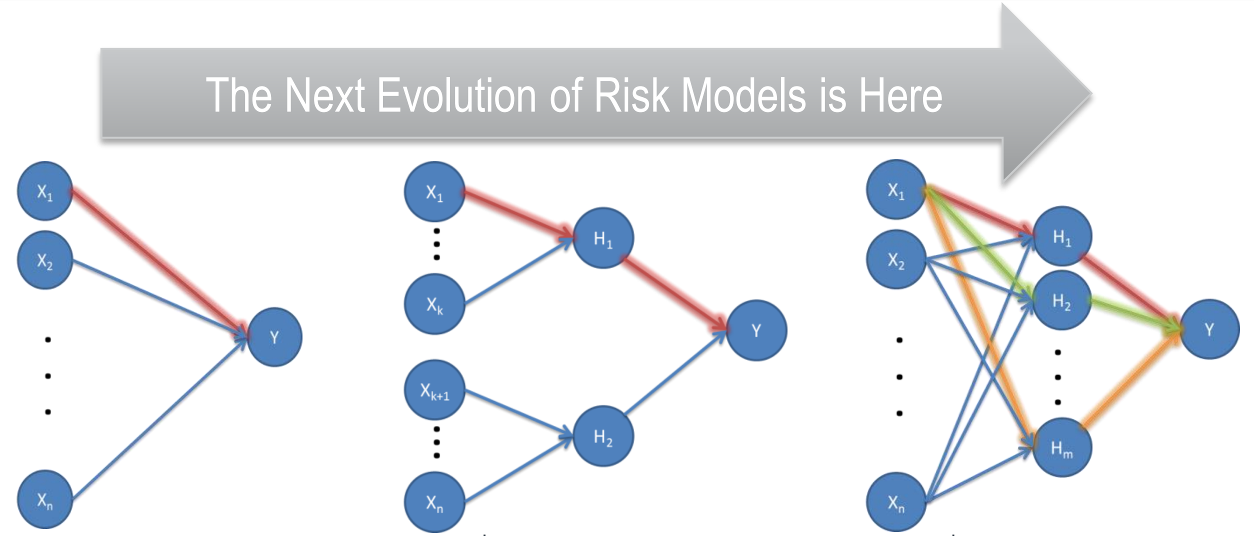
$$\begin{aligned} &f(x_1^m, \ldots, x_i^m, \ldots, x_n^m) - f(x_1^m, \ldots, X_i, \ldots, x_n^m) \\ &\quad = (\beta_0 + x_1^m\beta_1 + \cdots + x_i^m\beta_i + \cdots + x_n^m\beta_n) \\ &\qquad\qquad -(\beta_0 + x_1^m\beta_1 + \cdots + X_i\beta_i + \cdots + x_n^m\beta_n) \\ &\quad = \beta_i(x_i^m - X_i). \end{aligned}$$

THE CASE FOR NEURAL NETWORKS

NeuroDecision™ Technology – Superior Risk Assessment with Increased Customer Transparency

Features

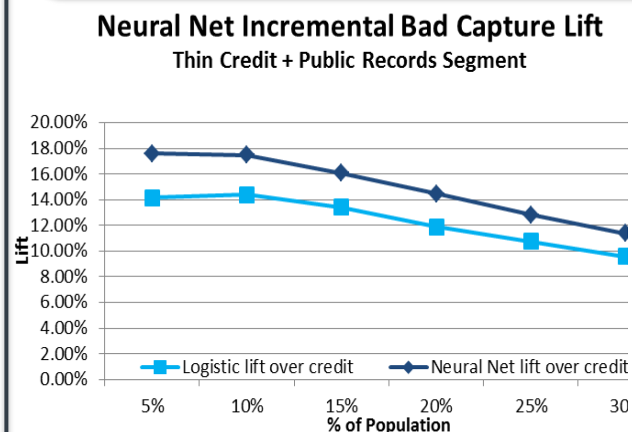
- Equifax exclusive patent pending machine learning technology producing regulatory compliant neural networks for risk decisioning applications
- Optimally constrained neural networks
  - Improved performance and accuracy
  - Returns a risk score and reason codes
  - Interpretable to customers and regulators
  - Enables deeper learning of consumer behavior through complex non-linear attribute interactions
- Applicable and deployable wherever traditional scorecards are appropriate
- Fully automated model evaluation expedites development in Cambrian



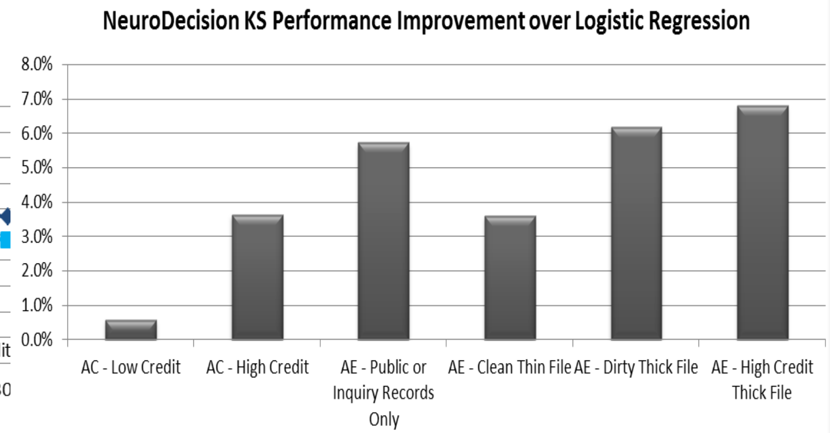
1st Generation: Logistic Regression

2nd Generation: Fusion and Embedded

3rd Generation: NeuroDecision

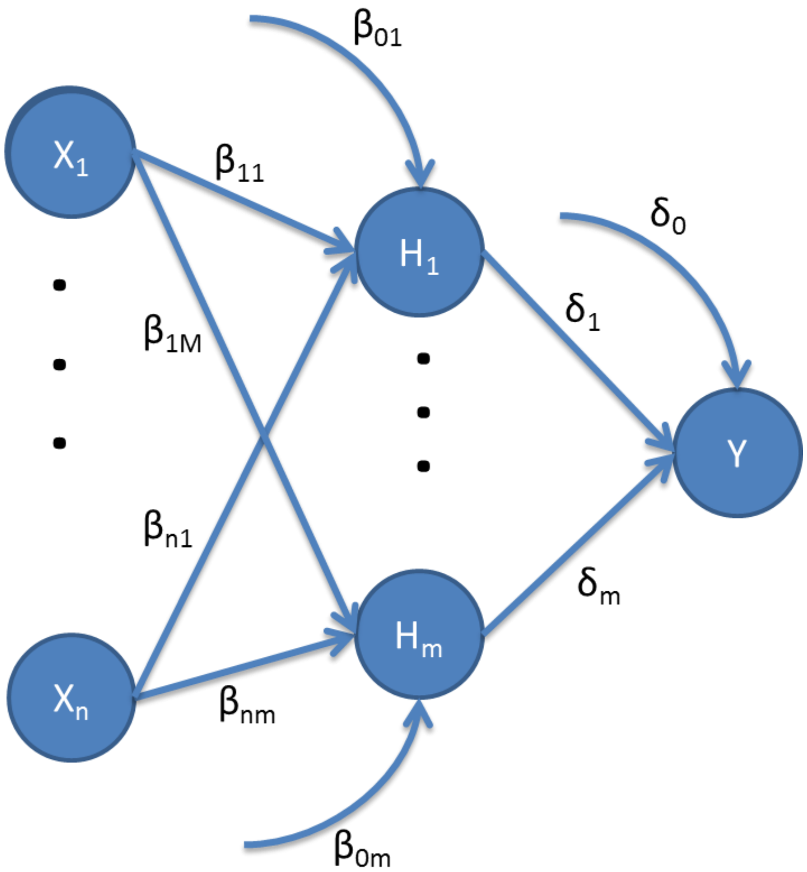


Custom Portfolio



Generic Portfolio

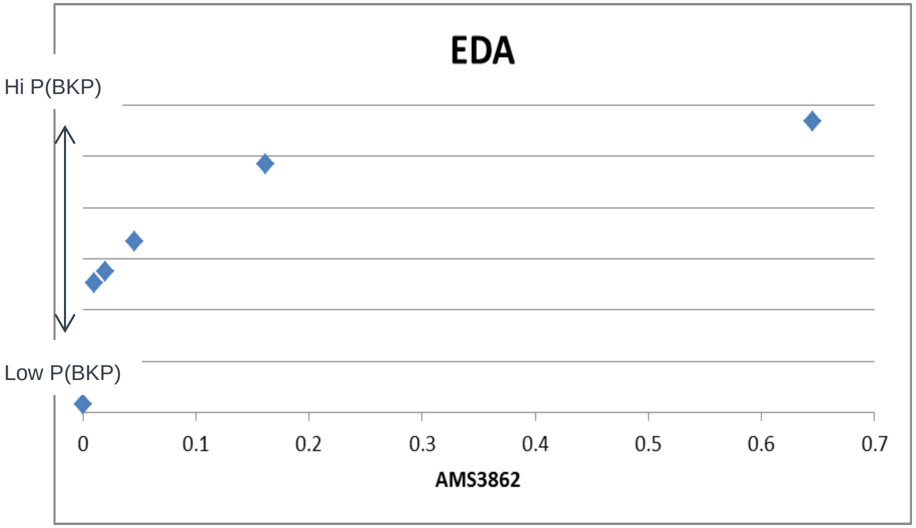
Neural Network Model

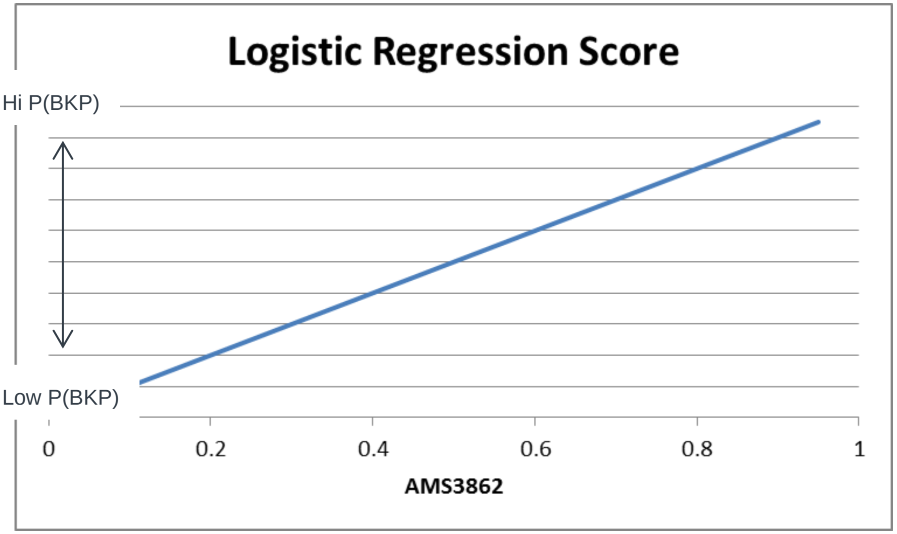


- $\beta_{ij}: X_i \to H_j$
- $\delta_j: H_j \to Y$
- $H_j = \frac{1}{1+\exp(-X\,\beta^j)}$
- $P(Y=1) = \frac{1}{1+\exp(-H\,\delta)}$
- $X = [1, X_1, \ldots, X_n]$
- $H = [1, H_1, \ldots, H_m]$
- $\beta_j = [\beta_{0j}, \beta_{1j}, \ldots, \beta_{nj}]^T$
- $\delta = [\delta_0, \delta_1, \ldots, \delta_m]^T$

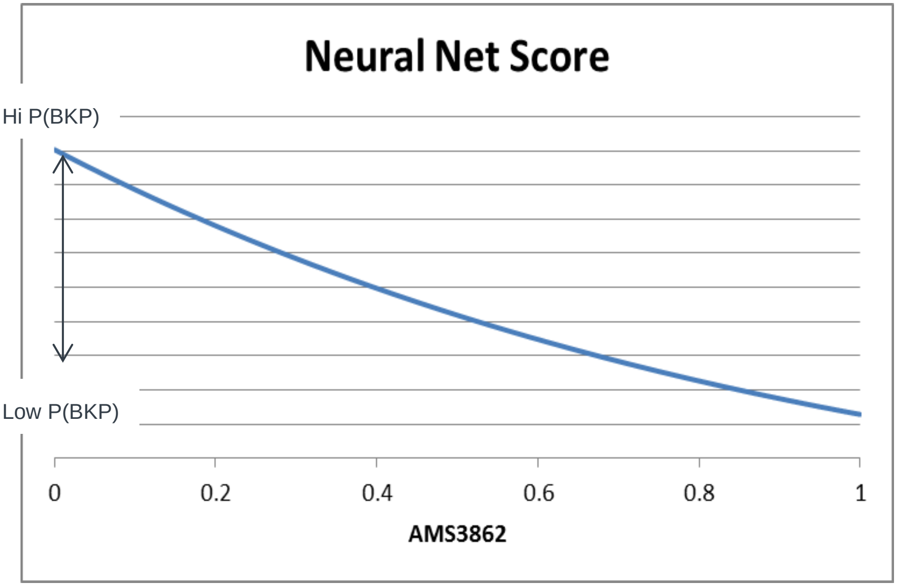
REGULATORY CHALLENGES WITH UNCONSTRAINED NEURAL NETWORKS

Challenge: Unconstrained Neural Networks May Violate Business Rules





- Odds of bankruptcy should increase as the percent of total past due amount to total balance on revolving trades (AMS3862) increases
- In an unconstrained Neural Net, **15%** of consumers' scores would actually **improve** if AMS3862 is increased



Challenge: The Change in Unconstrained Neural Network Scores Is Non-constant as Attribute Values Change

Logistic Regression

- Changes in an attribute impacts each consumer's score equally

$$\frac{\partial f}{\partial X_i} = \beta_i$$

- Positive changes in an attribute always produce either score increases or score decreases for every consumer

- Maximum partial score per attribute is equal for each consumer; maximum location as well

Unconstrained Neural Nets

- Changes in an attribute impacts each consumer's score variably

$$\frac{\partial}{\partial X_i}(H\ \delta) = \sum_{j=1}^{m} \beta_{ij}\delta_j \frac{\exp(-X\ \beta^j)}{\left(1+\exp(-X\ \beta^j)\right)^2}$$

- Positive changes in an attribute **may** produce score increases for some consumers but decreases for others.

- Maximum partial score per attribute varies by consumer; maximum location as well

NEURODECISION TECHNOLOGY

A Neural Network Framework Satisfying Regulatory and Business Requirements

NeuroDecision Modeling Process

1. Identify the set of candidate predictors satisfying relevant business rules and regulations. Perform exploratory data analysis (EDA) on this set. This includes analyzing the distribution of each predictor. The bivariate relationship of each predictor with the good-bad odds ratio is analyzed to quantify its predictive strength and trend.
2. Apply variable treatment to each predictor so that the observed bivariate relationship with the odds ratio is monotonic.
3. Apply preliminary variable reduction techniques to form a candidate set of treated predictors with monotonic relationships with the odds ratio.
4. Fit the single hidden layer artificial neural network model using industry and scientifically accepted standards.
5. Apply the *coefficient method*, an automated process to analyze $\frac{\partial}{\partial X_i}(H\ \delta) = \sum_{j=1}^{m} \beta_{ij}\delta_j \frac{\exp(-X\ \beta^j)}{(1+\exp(-X\ \beta^j))^2}$, to determine whether the modeled score P (Y = 1) exhibits a monotonic relationship with each predictor X_i for every consumer in the sample. For those predictors with a monotonic relationship with P (Y = 1), identify those whose trend is in accordance with the observed trend from step 1. These predictors are valid for NeuroDecision models. If any predictors are not valid, the developer must adjust the input predictors or the number of nodes in the hidden layer and return to step 1.
6. The process terminates when the modeled relationship of the score with each predictor is monotonic and the trend agrees with the EDA.

NeuroDecision Technology: *Constrained* Neural Net Model Development Parallels Traditional Scorecard Development

Logistic Regression

- Connection coefficient completely determines how a score reacts to attribute changes. Score changes are consistent per consumer
- Positive coefficient → Score Increase
- Visually compare coefficient to EDA
- Transform, bin, or discard attributes failing visual EDA comparison
- Rank order the differences $f(x_1^m, \dots, x_i^m, \dots, x_n^m) - f(x_1^m, \dots, X_i, \dots, x_n^m) = \beta_i(x_i^m - X_i)$

NeuroDecision Technology

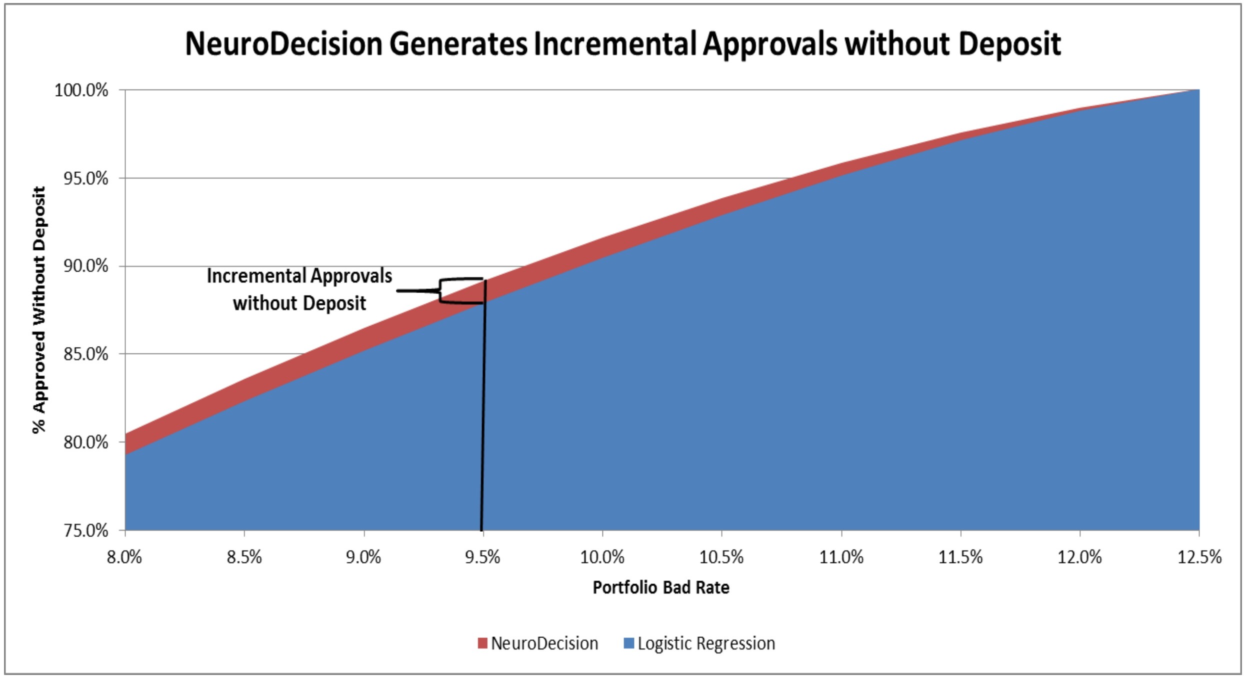
- "Coefficient method" aggregates changes per connection to determine how a score reacts to attribute changes. Score changes vary per consumer; direction is consistent with EDA
- Positive coefficient method → Score Increase
- Visually compare coefficient method to EDA
- Transform, bin, or discard attributes, or alter number of hidden nodes, for attributes failing visual EDA comparison
- Rank order the differences $f(x_1^m, \dots, x_i^m, \dots, x_n^m) - f(x_1^m, \dots, X_i, \dots, x_n^m)$

Note: Technical documentation of NeuroDecision Technology method is found in Appendix

Model Risk Management

- Model Attributes
  - Equifax business rules prohibit including any attribute describing a prohibited basis in risk models.
  - Correlation analysis is conducted on final model attributes that may serve as proxies for age.
  - No other prohibited bases are identifiable on the credit file.
- Model Scoring Domain
  - In traditional logistic regression, the scoring domain is assumed to be the Cartesian product of each attribute's domain.
  - In NeuroDecision models, the development sample domain may not sufficiently approximate the assumed Cartesian product domain.
  - Model Risk Management
    - The model is analyzed on the Cartesian product to determine if the reason code assignment is appropriate across the entire domain.
    - If not, the scoring domain is restricted to a convex subset (e.g. the convex hull of the development sample). Consumers exterior to the convex subset are uniquely mapped to the boundary.

Example: NeuroDecision Generates Lift in Both Tails of the Score Distribution



Cumulative Population	Logistic Regression	Neuro-Decision
100	100	100
90	98.28	98.43
80	95.51	95.68
70	91.64	92.06
60	86.73	87.23
50	80.38	80.89
40	71.93	72.58
30	60.99	61.64
20	45.87	47.37
10	26.32	28.21

- NeuroDecision generates higher approval rates for consumers while maintaining a constant risk portfolio profile
  - For a portfolio bad rate of 9.5%, 1.3% more consumers are approved
- NeuroDecision captures more bad consumers in the lower score range
  - Captures 3.3% more bad consumers in the bottom 20% of the population

Conclusions: NeuroDecision Technology Meets Criteria Described in Reg B and FCRA

- The modeling process is derived from existing business rules
- The business rules allow us to understand the complexities of the Neural Network
- The generalized coefficient method evaluates Neural Network risk models in the same way we evaluate Logistic Regression models. These models perform at least as well, and almost always outperform the predictive capability of logistic regression
- The output is a set of adverse action reason codes generated from the model, unique to each model, and appropriate given the consumer's credit risk behaviors
- The method for selecting reason codes works whether it is evaluating 'points below average' or 'maximum points lost', producing results that are substantially similar to methods described in the regulations

QUESTIONS?