PREDICITIVE MODELS WITH EXPLANATORY CONCEPTS

A GENERAL FRAMEWORK FOR EXPLAINING MACHINE LEARNING CREDIT RISK MODELS THAT SIMULTANEOUSLY INCREASES PREDICTIVE POWER

MATTHEW TURNER, PHD AND MICHAEL MCBURNETT, PHD

Equifax Data Science

Abstract. Lenders are required to transmit relatively few raw consumer credit behavior data values to credit reporting agencies. From the raw data, credit bureaus have derived thousands of predictive attributes. By construction, these derived attributes are highly correlated. Multicollinearity is known to hamper the ability to explain statistical and machine learning models. In general, it is desirable and often required to be able to explain the decision process of credit risk models. When a modeler attempts to incorporate numerous highly collinear attributes in a statistical model and maximize prediction, the impact of multicollinearity inhibits explainability. Traditional solutions to this problem include omitting variables with improperly signed parameters or large standard errors, or using factor analysis to extract a lower dimensional representation of the original variables prior to estimating model parameters. These solutions increase the ability to explain the model at the expense of decreasing its predictive power.

This paper describes a method of utilizing collinear attributes to increase the predictive power in a credit risk model while simultaneously generating model explanations. We make use of collinear attributes to construct a predictive model. The full predictive power of a modeling technique is realized by not reducing the dimension of the predictor space. Factor analysis is then used to generate interpretable concepts that are extracted from the predictor space and used for model explanations. The patent-pending processes embodied in Predictive Models with Explanatory Concepts can be utilized for machine learning or artificial intelligence transparency requirements (i.e. explainable AI) imposed by business decision makers or regulators.

Contents

E-mail address: `matthew.turner@equifax.com, michael.mcburnett@equifax.com`.
Date: August 8, 2019.

1. Introduction

Big data, cloud computing, machine learning, and artificial intelligence are all buzzwords in the analytics industry. But traditional credit scoring systems are not positioned to leverage the opportunities that big data enable. For example, logistic regression is a common technique used for credit scoring systems that does not take full advantage of the depth (number of rows) and breadth (number of available features) in the data. Traditional methods to produce credit scoring systems focus on explanatory models. Model development is accompanied by conservative restrictions on interrelationships in the data. Hence, the credit scoring industry has not kept pace with growth of technology and big data. How do we produce regulatory compliant and explainable machine learning models leveraging big data capabilities? In this paper we introduce Predictive Models with Explanatory Concepts (PMEC). It produces high dimensional models without multicollinearity and attribute significance constraints imposed on traditional regulatory compliant models.

Multicollinearity affects risk models on two fronts: model explainability and model fit. De Veaux and Ungar [1] showed that "neural networks tend to be fairly insensitive to problems of multicollinearity." Since neural nets are not adversely affected by modest amounts of multicollinearity, they can be fully leveraged as a credit scoring system if explainability can be achieved. PMEC was developed out of the desire to leverage collinear big data in a predictive model that can be explained.

1.1. Regulatory Requirements of Credit Scoring Systems.

Traditional logistic regression scorecard development procedures are entrenched in the credit scoring industry. Such models are well understood by end-users and regulators. Decisioning scorecards, such as risk scores, fall under the Fair Credit Reporting Act (FCRA) and Equal Credit Opportunity Act (ECOA) Regulation B. The Federal Reserve has also written interpretations of regulations. Some relevant regulations and remarks include:

(1) 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, . . . ;
(2) 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.
(3) Federal Reserve: "Regulation B neither requires nor endorses any particular method of credit analysis."[1]

One implication of item (1) is that key factors affecting the credit score should be both logical and actionable to the consumer. These key factors are typically referred to as *reason codes* or *adverse action codes* in the industry and identify the top 4 "reasons" an entity did not receive the best possible score. Logical and actionable reason codes in traditional logistic regression scorecard development are a direct result of established model development procedures intended to produce an explainable model. Such modeling techniques are staples among many sciences that focus on regression as a means to test hypotheses about causal relationships. This modeling methodology is known to construct an *explanatory model*.

[1] `http://www.federalreserve.gov/boarddocs/supmanual/cch/fair_lend_reg_b.pdf`

1.2. **Explanatory Models as Credit Scoring Systems.** In the credit scoring industry, explanatory models identify key variables that measure consumer behavior driving risk. The focus is to infer relationships between variables and outcomes seeking causal relationships. Explanatory models must be well-structured and interpretable. In explanatory models, attributes are used as explanatory variables and descriptions of the attributes are synonymous with reason codes.

The credit scoring industry has interpreted regulations to require explanatory modeling techniques and built logistic regression models accordingly. In explanatory models, there is little tolerance for interrelationships among the explanatory variables. Observational credit data attributes tend to be constructed such that a high degree of collinearity is present in many of them. This has the effect of inhibiting the explanatory power of these models.

1.3. **Predictive Models as Credit Scoring Systems?** At the other end of the modeling spectrum is *predictive models*. The goal with these models is to maximize predictive accuracy, which may sacrifice explainability. Such models seek mainly to predict future outcomes based on observed entity behavior. In predictive models, model specification is de-emphasized, while hold-out and out of time sample validation is required to assess model stability. Interrelationships are not directly assessed and no emphasis is placed on the explanatory nature of the model. As a result, reason codes are not applicable to purely predictive models and do not satisfy regulatory requirements.

1.4. **Multicollinearity in Credit Scoring Systems.** Interrelationships are routinely described and measured in terms of *multicollinearity*. Multicollinearity is the condition observed when one explanatory variable is a linear combination of one or more other explanatory variables. Formally, there exists constants a_i, not all 0, such that $\sum_{i=1}^{n} a_i X_i = 0$. Multicollinearity is also used to describe the case when $\sum_{i=1}^{n} a_i X_i \approx 0$. Multicollinearity is usually measured in terms of the Variance Inflation Factor (VIF) available in most statistical software packages. Industry standards for allowable VIF values in risk models range from 2 for conservative customers to values of 4-5 in the most extreme instances. There are known numerical and statistical problems when multicollinearity is present in the data. These include

- unstable parameter estimates in explanatory models, including the possibility of an "incorrect" or illogically signed parameter that is opposite the observed trend in the data,
- large standard errors of parameter estimates,
- and the numeric stability of required design matrix inverse calculations.

In the credit scoring industry, there is another more prevalent problem with multicollinearity: reason code assignment. Current industry development standards limit a credit scoring system's predictive ability by leveraging antiquated reason code methods. Though these traditional methods are regulatory compliant, they are burdensome in their conservative approach to compliance. First, by minimizing VIF and restricting interrelationships among attributes in a model, the assumption is credit attributes move independently of one another. Of course, this is an unreasonable assumption unless the attribute space is greatly reduced. Second, reason code methods rely on attributes that move independently of one another. Other than attributes such as *Months on File*, most attributes are at least moderately correlated with multiple other attributes. By computing reason codes under the assumption that attributes change independently of one another, true underlying factors driving score changes are diminished in importance while uncorrelated attributes such as *Months on File* are improperly emphasized in their importance. This approach assumes the scoring domain is the Cartesian product of the input attributes. Even in the most conservative of explanatory models with VIFs smaller than 2, one finds that the observed scoring domain is a strict subset of the theoretical Cartesian product scoring domain. It is worth reemphasizing that such models often leave substantial performance gains on the table.

For the purpose of regulatory compliance, explanatory models are quite adequate. In particular,

(1) Models are readily interpretable, by examining the signed impact of an attribute on the credit score, for causal inference, and
(2) Reason codes are readily returned since conservative VIF requirements ensure the assumption that attributes move independently is not unreasonable.

However, we contend that such models do not best serve credit granting institutions nor persons seeking credit. Explanatory modeling techniques may lead to variables being dropped for the wrong sign, large VIF, large standard error, or insignificant parameter estimates even though the model fit may be superb and prediction stable. Credit scoring systems with more predictive models allow lenders to better manage risk by more accurately assessing risk and benefit credit seeking consumers by allowing for more accurate risk-based pricing.

1.5. **Academic solutions to Multicollinearity.** How is multicollinearity handled in most modeling applications? Standard remedies are (e.g. Hanushek and Jackson [2]):

- Find another sample with uncorrelated variables
- Find instruments for correlated variables
- Remove variables through some selection process (e.g. varclus)
- Conduct a regression using factor analysis or principal components as predictors
- Run a constrained regression, such as ridge regression

In general, none of these solutions helps in risk modeling. Part of the problem is that, in general, we do not have a theory. If we had a theory, we would know the equations governing risk. Even if we had a theory, some attributes known to be highly predictive are forbidden from use. These attributes are found in the disturbance term and likely correlated with other predictors. Hence, not only must we contend with biased estimates we also are forced to mine the data for the best predictive model that can be explained.

1.6. **Predictive Models with Explanatory Concepts (PMEC).** In light of the points made by the Federal Reserve Board that "Regulation B neither requires nor endorses any particular method of credit analysis" previously discussed in section (1.1) bullet (3), we develop a novel solution to the collinearity problem in risk models. Our solution to multicollinearity is inspired by Leahy [3] and his commentary on stable sums of coefficients. While academic solutions tend to focus on the accuracy and interpretability of estimates in explanatory models and on stable coefficients over time in predictive models, in the realm of risk models we need to consider both. Predictive models are better for credit scoring but we modify model specifications to ensure we develop an Explanatory Model that is more predictive and satisfies regulatory requirements.

PMEC finds components that move (mostly) independently and returns reason codes that are the real drivers of score differences. It finds the underlying components that move independently (or mostly independent in the case of oblique rotations) and quantifies their impact on the credit score. This provides a much better picture to the consumer on what actions should be taken and theoretically is more sound than current practice in which we pretend the scoring domain is Cartesian and attributes move independently.

In many cases, data used to predict credit risk is collinear by construction. Consider two attributes: Number of times the consumer has been 30 days past due in the last six months, and number of times the consumer has been 60 days past due in the last six months. Every consumer who is 60 days past due in the last six months had to pass through the 30 days past due attribute to get there. In general, not every consumer who is 60 days past due is also 30 days past due in the last six months. This means these two attributes are highly correlated rather than perfectly collinear. Hence, both can be used in a model, but the degree of collinearity may impact the variance of the estimates, or induce one of these attributes to have the wrong sign in a multivariate model. If we could use both attributes in a single model it would be more predictive; if one attribute has the wrong sign it is not explainable. We either have to delete it from the final model or risk not being able to return rational, logical reason codes that can help the consumer improve his or her score.

Figure 1. *Predictive Models with Explanatory Concepts* bridges the gap between predictive and explanatory models to produce best in market regulatory compliant risk scores.

Predictive Models with Explanatory Concepts bridges the gap between these two extremes to produce best in market regulatory compliant risk scores. PMEC is a mathematical model development framework that simultaneously unifies Equifax's current logistic regression and neural network technology. It works by constructing a factor analysis of the correlated attributes in the model, and using those factors as the basis for reason codes. The reason codes can be used to fully explain the model and consumers whom act rationally on the information returned will always improve their credit score. Predictive Models with Explanatory Concepts is a proprietary algorithm that produces

- An advanced predictive model coupled with
- A Factor Analysis of the input variables to produce
- Concepts fully explaining the predictive models behavior.

2. PMEC

Let $f(X_1, \ldots, X_n)$ be a risk indicator, such as a credit score, built using, for example, logistic regression or two hidden layer feed forward neural network modeling technique. As our main goal is rank-ordering risk, and since monotonic transformations retain rank-ordering, we may consider f to be on a "log-odds" scale. Then

$$f(X_1, \ldots, X_n) = \begin{cases} X\beta, & \text{logistic regression,} \\ H^2\delta, & \text{neural networks,} \end{cases} \tag{1}$$

where φ is a differentiable sigmoid activation function (such as logistic function) and

$$H_k^2 = \varphi\left(H^1\beta_{\cdot k}^2\right), H_j^1 = \varphi\left(Z\beta_{\cdot j}^1\right), Z_i = \frac{X_i - \mu_i}{\sigma_i}. \tag{2}$$

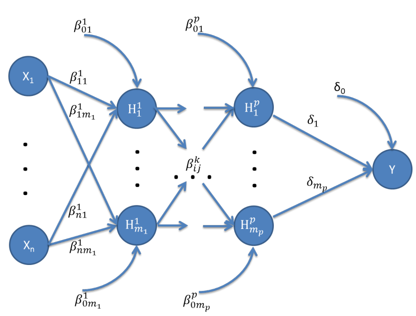

Figure 2. Multiple hidden layer feed-forward neural network.

When multicollinearity exists, traditional remedies include applying principal components analysis (PCA) or Factor Analysis (FA) to reduce the dimension of the data into a set of orthogonal or less correlated variables called components or factors. Factor analysis in particular, is a statistical method to explain the correlation structure among variables while aiming to produce uncorrelated interpretable factors. Traditional methods regress on the newly created components or factors that are significant (see e.g. figure (2)).

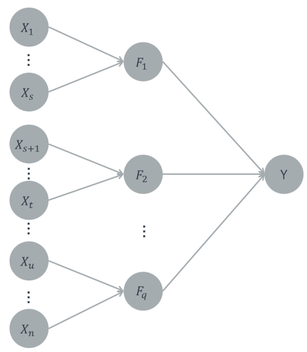

Figure 3. Traditional academic solution to Multicollinearity.

PMEC solves a separate business problem. We are not concerned with multicollinearity as long as prediction is stable and accurate. In fact, this work is motivated by two observations. First, multicollinearity is a feature that can be exploited in machine learning techniques such as neural networks as "neural networks tend to be fairly insensitive to problems of multicollinearity." (De Veaux and Ungar [1]). Second, in a short but well written article, Leahy [3] observed that the sum of standardized coefficients on collinear attributes tends to remain stable. This observation is profound. Before discussing, consider the following simple experiment to illustrate this observation of multicollinearity. Five trials were run on simulated data. In each trial, 5000 rows of two highly correlated variables were simulated, along with a dependent variable $Y \sim \mathit{Bernoulli}\left(1/\left(1+\exp\left(0.5X_1+0.5X_2\right)\right)\right)$. A logistic regression model was then fit to the data with the results provided in table (1).

Trial	β_1	β_2	Standardized β_1	Standardized β_2	Standardized $\beta_1+\beta_2$
1	0.1023	0.7869	0.0163	0.131	0.1473
2	-0.4508	1.3301	-0.0722	0.2224	0.1502
3	0.6385	0.3637	0.1013	0.0606	0.1619
4	0.4338	0.7016	0.0692	0.1174	0.1866
5	0.459	0.5617	0.0728	0.0924	0.1652

Table 1. The sum of standardized coefficients on multicollinear attributes tends to remain stable.

Table (1) provides the standardized coefficients of the logistic regression model of each trial. The estimates are very unstable as expected. In each trial, the standard error was ≈ 0.34 and hence, the estimates are not statistically significant from zero. In fact, we can see in trial 2 that β_1 actually flips signs and is negative. However, we see in column 3 that the sum of the standardized coefficients are stable, ranging from 0.1473 to 0.1866. PMEC develops a formal theory leveraging the preceding two facts.

To solve this problem, we use factor analysis to explain significant driving factors in a predictive model. Consider first the factor analysis model

$$Z_i = \frac{X_i - \mu_i}{\sigma_i} = LF + \varepsilon_i = \sum_{j=1}^{q} \ell_{ij} F_j + \varepsilon_i. \tag{3}$$

Factor Analysis is a latent variable model in that the factors F_j are not directly observable and measured. To regress on F, approximation techniques must be employed to estimate each F_j from the observed data X. As well, F does not account for all variance in each X. Fitting on X has been observed to produce more predictive models. The traditional academic concern is the standard errors of the coefficients and the stability of coefficients across samples. This limits the explanatory power of the model. However, the goal of a predictive model is not explanatory power but predictive power. To satisfy regulation, key factors affecting the credit score must be provided to the consumer or entity. PMEC accomplishes this by building a predictive model on (possibly) collinear big data and applies common factors to explain the key factors of the model as depicted in figure (2).

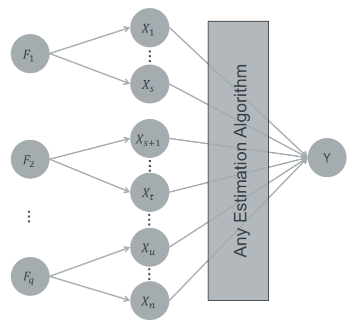

FIGURE 4. Predictive Model with Explanatory Concepts

We require

(1) The VIFs between $F_1, \ldots, F_q, \varepsilon_1, \ldots, \varepsilon_n$ to be small. According to industry convention, we will assume that each of the factors move independent of one another, so that variable effects can be considered component wise. This is a reasonable assumption since the theory of Factor Analysis assumes each are independent and are rotated by orthogonal transformation. Care must be taken when considering oblique rotations.
(2) The trend of each common factors F_j in the model must be correct (discussed in sections (2.1) and (2.2)).
(3) The Factor Analysis fit must conform to academic standards and the common factors must be interpretable.

Motivated by the fact that for a good factor fit, $var(\varepsilon_i) = 1 - \sum_{j=1}^{q} \ell_{ij}^2 << 1$ and that f is continuous and sufficiently smooth in its inputs, let us rewrite the score as

$$
\begin{aligned}
f(X_1, \ldots, X_n) &= f\left(\mathbb{E}\left(X_1|F_1, \ldots, F_q\right), \ldots, \mathbb{E}\left(X_n|F_1, \ldots, F_q\right)\right) \\
&\quad + \left(f(X_1, \ldots, X_n) - f\left(\mathbb{E}\left(X_1|F_1, \ldots, F_q\right), \ldots, \mathbb{E}\left(X_n|F_1, \ldots, F_q\right)\right)\right) \\
&\stackrel{def}{=} f\left(\mathbb{E}\left(X_1|F_1, \ldots, F_q\right), \ldots, \mathbb{E}\left(X_n|F_1, \ldots, F_q\right)\right) + \varepsilon, \quad (4)
\end{aligned}
$$

where ε is defined by

$$
\varepsilon \stackrel{def}{=} f(X_1, \ldots, X_n) - f\left(\mathbb{E}\left(X_1|F_1, \ldots, F_q\right), \ldots, \mathbb{E}\left(X_n|F_1, \ldots, F_q\right)\right).
$$

The first term in equation (4) is the component of the score driven entirely by the common factors. This component provides a much better indicator of what is driving a score since $\varepsilon \approx 0$. Since credit attributes tend to be collinear by construction, traditional explanatory models indicate that attributes that are most orthogonal to other predictors (e.g. *age of oldest account*) are the key factors of the model and are pushed to the top of the rankings. The effect of underlying concepts that may have several related attributes in the model, such as delinquency, are driven down in importance. Each variable that measures the concept of delinquency is important to the risk score (assuming each is significant in the model), and contain subtle distinct information that improves the accuracy of the risk score.

2.1. **PMEC for Logistic Regression.** Let us first apply the PMEC model depicted in figure (2) to logistic regression and determine how the common factors impact the credit score. The model is given by (working on the log-odds scale)

$$f(X_1, \dots, X_n) = X\beta, Z_i = \frac{X_i - \mu_i}{\sigma_i} = LF + \varepsilon_i = \sum_{j=1}^{q} \ell_{ij} F_j + \varepsilon_i. \tag{5}$$

To determine how the common factor F_p influences the model, we consider the partial derivative

$$\frac{\partial f}{\partial F_p} = \sum_i \frac{\partial f}{\partial X_i} \frac{\partial X_i}{\partial Z_i} \frac{\partial Z_i}{\partial F_p} = \sum_i \beta_i \sigma_i \ell_{ip}. \tag{6}$$

If F_p is represents positive behavior (such as satisfactory payments), we require $\sum_i \beta_i \sigma_i \ell_{ip} \geq 0$. If F_p represents negative behavior (such as utilization), we require $\sum_i \beta_i \sigma_i \ell_{ip} \leq 0$. Let us make an important observation about equation (6). In a traditional logistic regression explanatory model, the sign of the coefficient of each variable is interpreted and analyzed against data and business knowledge. Variables with counter-intuitive or unexplainable signs are removed from the model. But for PMEC, only the aggregate affect of all variables that load on a common factor is important. In particular, individual attributes may have an improper sign while the common factor behaves properly in the model. This flexibility allows for multicollinear variables to be included in the model.

2.2. **PMEC for NeuroDecision.** We introduced NeuroDecision® at the Credit Scoring and Credit Control XIV in 2015. NeuroDecision is our patented explainable Artificial Intelligence solution that optimally constrains neural networks to be monotonic and generate key factors driving the model. PMEC can be applied to the neural networks to relax the monotonicity constraints, allowing for an even more predictive model that can generate key factors driving the model. The model is given by (working on the log-odds scale)

$$Y = H^2\delta, H_k^2 = \varphi\left(H^1 \beta_{\cdot k}^2\right), H_j^1 = \varphi\left(Z\beta_{\cdot j}^1\right), Z_i = \frac{X_i - \mu_i}{\sigma_i} = L_{i\cdot}F + \varepsilon_i. \tag{7}$$

To determine how the common factor F_p influences the model, we again consider the partial derivative

$$\begin{aligned}
\frac{\partial Y}{\partial F_p} &= \sum_k \frac{\partial Y}{\partial H_k^2} \frac{\partial H_k^2}{\partial F_p} \\
&= \sum_k \frac{\partial Y}{\partial H_k^2} \sum_j \frac{\partial H_k^2}{\partial H_j^1} \frac{\partial H_j^1}{\partial F_p} \\
&= \sum_k \frac{\partial Y}{\partial H_k^2} \sum_j \frac{\partial H_k^2}{\partial H_j^1} \sum_i \frac{\partial H_j^1}{\partial Z_i} \frac{\partial Z_i}{\partial F_p} \\
&= \sum_k \sum_j \sum_i \frac{\partial Y}{\partial H_k^2} \frac{\partial H_k^2}{\partial H_j^1} \frac{\partial H_j^1}{\partial Z_i} \frac{\partial Z_i}{\partial F_p} \\
&= \sum_k \sum_j \sum_i \delta_k \beta_{jk}^2 \beta_{ij}^1 \ell_{ip} \varphi'\left(H^1 \beta_{\cdot k}^2\right) \varphi'\left(Z\beta_{\cdot j}^1\right) \\
&= \sum_k \sum_j \varphi'\left(H^1 \beta_{\cdot k}^2\right) \varphi'\left(Z\beta_{\cdot j}^1\right) \left(\sum_i \ell_{ip} \beta_{ij}^1 \beta_{jk}^2 \delta_k\right).
\end{aligned} \tag{8}$$

Since $\varphi'(\cdot) > 0$, a sufficient condition to ensure positive monotonicity is:

$$\sum_i \ell_{ip}\beta^1_{ij}\beta^2_{jk}\delta_k \geq 0 \quad \forall \quad j,k. \tag{9}$$

Of course a similar condition holds for negative monotonicity. The appropriate trend is enforced during model development to ensure that positive factors have a positive trend and negative factors have a negative trend. Recall that in the logistic regression model in the preceding section (2.1), only the aggregate affect of all attributes that load on a factor are important to ensure the factor behaves properly in the model. We now see a similar situation for neural networks. The sufficient condition that we enforce given by equation (9) shows that, although individual attributes in the model may not have strictly monotonic trends, the factor has a strictly monotonic trend that is controlled by the aggregate effect of all the attributes that load on a factor in the neural network. This allows us to greatly relax the attribute monotonicity constraints of the original NeuroDecision algorithm. Now, the trends of individual attributes can be non-monotonic as long as the factor is monotonic.

3. Key Factors of a PMEC Credit Scoring System

Now we can apply any standard reason code technique to compute the impact of factor F_p on the credit score. For example, equation (10) generalizes a standard "points below max" approach widely used for logistic regression credit scoring systems:

$$g(F_1,\ldots,F_p^*,\ldots,F_q,\varepsilon_1,\ldots,\varepsilon_n) - g(F_1,\ldots,F_p,\ldots F_q,\varepsilon_1,\ldots,\varepsilon_n). \tag{10}$$

Since the factors are unobservable latent variables, they must be estimated. Denote the estimates, called factor scores, by $\hat{F}_p$ and $\hat{\varepsilon}_i$. There are various techniques for estimating the factor scores $\hat{F}_p$, but most importantly the factor scores are linear combinations of the input X attributes. Of course, $\hat{\varepsilon}_i = \hat{\varepsilon}_i|_{\hat{F}_p} = Z_i - L\hat{F}$. For purposes of explaining the model, we will not actually need to compute the factor scores as explained below. Let $\hat{F}_p^*$ denote the location of the factor score $\hat{F}_p$ that maximizes the score $g(\hat{F}_1,\ldots,\hat{F}_q,\hat{\varepsilon}_1,\ldots,\hat{\varepsilon}_n)$. Since g is monotonic in the common factors, $\hat{F}_p^*$ will be the right or left endpoint of the domain of $\hat{F}_p$, depending on whether it is a positive or negative behavior factor. Owing to the fact that the factor scores $\hat{F}_p$ are linear inputs of the input X attributes, $\hat{F}_p^*$ will correspond to a right or left endpoint of each X_i that loads on F_p. Therefore, we have $X_i = X_i^*$ and $\hat{\varepsilon}_i|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q} = Z_i^* - \sum_{k\neq p}\ell_{ik}\hat{F}_k - \ell_{ip}\hat{F}_p^*$ for every attribute X_i that loads on the factor $\hat{F}_p$. Thus at $\hat{F}_p = \hat{F}_p^*$, we have $Z_i = L\hat{F} + \hat{\varepsilon}_i = \sum_{k\neq p}\ell_{ik}\hat{F}_k + \ell_{ip}\hat{F}_p^* + Z_i^* - \sum_{k\neq p}\ell_{ik}\hat{F}_k - \ell_{ip}\hat{F}_p^* = Z_i^*$. Now we can apply the "points below max" equation (10) to get

$$\begin{aligned} &g(\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q,\hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q},\ldots,\hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q}) \\ &\qquad - g(\hat{F}_1,\ldots,\hat{F}_p,\ldots\hat{F}_q,\hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q},\ldots,\hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q}). \end{aligned} \tag{11}$$

Equation (11) is the formula that we apply at Equifax to generate key factors (a.k.a. reason codes or adverse action codes) of a PMEC model. In the next section (3.1), we will look at some specific examples and show how the PMEC reason code formula generalizes the traditional "points below max" key factor approach.

3.1. Examples.

Let us turn to some specific examples. First, let's consider f as a logistic regression model and assume that the common factor F_p is a trivial factor with only X_i loading on F_p.

Applying equation (11) gives

$$\begin{aligned}
g(\hat{F}_1, \ldots, \hat{F}_p^*, \ldots, \hat{F}_q, \hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q}, \ldots, \hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q}) \\
- g(\hat{F}_1, \ldots, \hat{F}_p, \ldots \hat{F}_q, \hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q}, \ldots, \hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q}) \\
= f(X_1, \ldots, X_i^*, \ldots, X_n) - f(X_1, \ldots, X_i, \ldots, X_n) \\
= \beta_i(X_i^* - X_i). \quad (12)
\end{aligned}$$

Anyone applying "points below max" previously will quickly recognize equation (12) as the standard formula applied to compute key factors for a logistic regression model. One simply takes the points that the consumer could have attained for the input X_i minus the points they actually received. So, in this case, equation (11) generalizes the conventional "points below max" for logistic regression.

As a second example, let f continue to be a logistic regression model but assume that X_r, X_s, and X_t are the variables that load on the common factor F_p. Applying equation (11) gives

$$\begin{aligned}
g(\hat{F}_1, \ldots, \hat{F}_p^*, \ldots, \hat{F}_q, \hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q}, \ldots, \hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q}) \\
- g(\hat{F}_1, \ldots, \hat{F}_p, \ldots \hat{F}_q, \hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q}, \ldots, \hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q}) \\
= f(X_1, \ldots, X_r^*, \ldots, X_s^*, \ldots, X_t^*, \ldots, X_n) - f(X_1, \ldots, X_r, \ldots, X_s, \ldots, X_t, \ldots, X_n) \\
= \beta_r(X_r^* - X_r) + \beta_s(X_s^* - X_s) + \beta_t(X_t^* - X_t). \quad (13)
\end{aligned}$$

In this case, we see that the "points below max" simply add together. For each input attribute that loads on a factor, we compute the difference between the maximum points a consumer could have attained for that attribute and the points the consumer actually attained and sum these differences. Again, this is a common approach and is the specification in the Predictive Model Markup Language (PMML). Even if we were not working on the log-odds scale, the rank ordering of the key factors remains the same since the logistic transformation is monotonic. Therefore, we can compute the key factors by adding the points lost together for each input attribute that loads on a factor.

We illustrate these two examples to emphasize that PMEC, when applied to logistic regression, produces an industry standard credit scoring system complete with key factors. PMEC though is more general and can be applied to any monotonic machine learning technique, including our monotonic neural network technique NeuroDecision. In the case that F_p is a trivial factor, the key factor equation (11) is

$$\begin{aligned}
g(\hat{F}_1, \ldots, \hat{F}_p^*, \ldots, \hat{F}_q, \hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q}, \ldots, \hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q}) \\
- g(\hat{F}_1, \ldots, \hat{F}_p, \ldots \hat{F}_q, \hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q}, \ldots, \hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q}) \\
= f(X_1, \ldots, X_i^*, \ldots, X_n) - f(X_1, \ldots, X_i, \ldots, X_n). \quad (14)
\end{aligned}$$

Equation (14) no longer has the convenient simplification of equation (12) that we get for logistic regression, since we no longer are working with a linear model. However, it is still easily applied to compute how a consumer's score could increase if they were at the optimal value X_i^* of X_i.

As a final example, let's apply the PMEC key factor equation (11) to a NeuroDecision model with three input variables X_r, X_s, and X_t loading on the factor F_p. In this case, equation (11) becomes

$$\begin{aligned}
g(\hat{F}_1, \ldots, \hat{F}_p^*, \ldots, \hat{F}_q, \hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q}, \ldots, \hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p^*,\ldots,\hat{F}_q}) \\
- g(\hat{F}_1, \ldots, \hat{F}_p, \ldots \hat{F}_q, \hat{\varepsilon}_1|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q}, \ldots, \hat{\varepsilon}_n|_{\hat{F}_1,\ldots,\hat{F}_p,\ldots,\hat{F}_q}) \\
= f(X_1, \ldots, X_r^*, \ldots, X_s^*, \ldots, X_t^*, \ldots, X_n) - f(X_1, \ldots, X_r, \ldots, X_s, \ldots, X_t, \ldots, X_n). \quad (15)
\end{aligned}$$

As in the preceding case, we don't get a convenient simplification as we did in logistic regression, since we no longer are working with a linear model. Also, the points lost no longer simply add together. But the equation (15) is still straight-forward to apply and determine the impact of the factor F_p on the model. Moreover, the PMEC key factor equation (15) accounts for the fact that, for multicollinear attributes X_r, X_s, and X_t, these attributes can't move independent of one another and must move together. This is powerful in that it does not favor input attributes that are orthogonal to the rest of the data and provides a much better explanation of the key factors impacting a credit score.

4. Conclusion

PMEC is designed to leverage big data to maximize predictive power while explaining key factors driving the outcome by striking a proper balance between predictive and explanatory models. The predictive power comes from leveraging big data in machine learning algorithms while Factors are use to explain the key factors driving the model. Factor Analysis is leveraged because the goal of Factor Analysis is to produce interpretable latent variables. The model developer can focus on predictive power by modeling directly on the big data X instead of a reduced dimension of Principal Components Analysis or Factor Analysis variables. During model development, the model developer need only focus on stable predictions and explainable factors. The advantage is that now big data, including multicollinear variables, can be leveraged in any relevant credit scoring system.

References

[1] De Veaux, R. D. and Ungar, L. H. (1994). *Multicollinearity: A tale of two nonparametric regressions*, pages 393–402. Springer New York, New York, NY.

[2] Hanushek, E. A. and Jackson, J. E. (1977). *Statistical methods for social scientists*. Academic Press New York.

[3] Leahy, K. (2001). Multicollinearity: When the solution is the problem. In Rud, O. P., editor, *Data Mining, Cookbook: Modeling Data for Marketing, Risk and Customer Relationship Management*. John Wiley & Sons, Inc., New York, NY, USA.