Determining the Factors Associated with Second Job Position: Using Adjacent Category Logit Model

K. Nirmalkanna[*1], E. R. A. D. Bandara[2] and L. C. S. Wijayasinghe[3]

University of Colombo, Sri Lanka.
Corresponding Author: knkanna87@gmail.com

Received: September 3, 2013 / Revised: February 24,2014 / Accepted: March 14, 2014

ABSTRACT

This study focuses on the factors that influence the second job position of an individual. The Corresponding response variable has four levels such as senior management, middle management, junior management and non executive grade which is in ordinal scale. Rather than sticking into classical approaches, a rarely used approach, adjacent category logit model approach has been used to accomplish this task. Logit structure of this model is most appropriate in this context. The predicting ability of the fitted model has been justified using artificial neural network approach. Final model has 3 separate equations for 3 logits and each provides separate information about one step increments in second job position. The covariate, organization of second job is highly influencing all 3 logits, individual's university and the sector of first job are influencing partially. Analysis reveals that these 3 factors are highly significant on influencing the response variable.

Keywords: Artificial Neural Network, Multi-categorical Response Variable, Multinomial Distribution, Hyperbolic Tangent, Softmax Activation Function, ROC Curve.

1. Introduction

Today finding a desired and satisfied job is not an easy task. Obtaining a higher position is much trickier. Therefore, many job seekers tend to use various career marketing tactics to become winning candidate for their potential employers. Many people suffer from not getting their desired job position. They believe that they have already satisfied the required qualifications for their target career. Therefore, many employees leave their current job, when they find their preferred job position in another company. However, this is a disadvantage for the organizations, since the allocated cost for the trainings of those employees is no more useful to them and also hiring new employees again is an additional cost.

In employer's point of view, if he or she has an appropriate knowledge about the current job market then they can avoid these problems by providing suitable concessions to the employees. In employee's point of view, if he or she thinks that it is not the desired and satisfied place to work in, then he or she needs appropriate knowledge about the current job market to move on next company. Eventually these two scenarios need the same information. An appropriate statistical approach has been used in this study to reveal such information from an existing data set.

The data set was obtained from the National Science Foundation of Sri Lanka. It consists of the information about employees who graduated in Sri Lankan universities. The data set contains 16 variables and all are categorical variables such as employee's gender, degree, university, information about the second job, information about the first job and employee's career objectives. This study only focuses on the individuals, who has a second job in management field, to narrow down the study. Since there were many missing and inappropriate values, a suitable cleaning process was applied, and finally a data set with 291 observations was obtained.

Adjacent category logit (ACL) model approach has been used to accomplish the task of determining the factors associated with the second job position, and the result has been justified using a neural network approach. This study discusses both of these approaches in detail in separate chapters. The response variable of interest is the second job position which consists of four categories such as senior management, middle management, junior management and non executive grade, and this variable is in ordinal scale. Both approaches give the result that the sector of the first job and the organization of the second job are highly influenced the second job position.

The neural network approach has been carried out using SPSS software. A best neural net with highest accuracy has been constructed. Further, it can be used to make predictions of a new instance. The procedure of both approaches, the findings together with a comparison, limitations of the study and further improvements have been discussed in a later section.

2. Literature Review

This study mainly focused the factors that affect the second job position of an employee, who is currently engaged in work, but seeking for a managerial position in another company. Reasons for this condition may be due to the unsatisfactory of salary or job-environment or for higher positions. Many studies have identified the factors that affect the job satisfaction (Ganzach, 1998;

Koh and Boo, 2001; Volkwein and Zhou, 2003; Wharton et al., 2000). However, up to our knowledge, no studies were conducted to identify the factors affecting next or second job position. Therefore, this is a new way of approach to solve such problems.

Our response variable is the second job position which is an ordinal categorical variable. Usually people try to analyse this kind of data with some commonly used models such as a base-line category logits model, proportional odds model and continuation-ratio logits model. But there exist another model called ACL model. Only two studies (Agresti, 1992; Goodman, 1983) were found for ACL model. It implies that this is very rarely used model. In Agresti's (1992) study, he fitted a very simple ACL model but in Goodman's (1983) study, he compared some log-odd models, ACL model was the one of them.

In our problem base-line category logits can explain moving to immediate higher position as well as to every other higher level. Therefore, it is only applicable when the current position is the lowest position. However, the proportional odds model illustrates group improvements, considering current and lower levels as one group while all higher levels as another group. Similarly, continuation-ratio logits consider current level and the group of all higher levels. Finally adjacent categories logits demonstrate an idea about moving to immediate higher level. The main advantage of ACL model is the corresponding logits giving more precise information about the levels of response variable.

An employee always looks for an immediate higher position from the current one, to reach his or her personal goals. It is not easy to jump to a position which is higher in several levels. For an example, senior management position as the next higher position cannot be expected from a person who is in the current position as a non executive grade. Hence, it is very meaningful, if it can be modeled for step by step improvement in the positions. Therefore, ACL model is the only model that can illustrate the way of step by step improvement.

Clearly, with a rarely used model approach, this is a new way to solve the problems which are mentioned in the introduction part. According to studies of Tam and Kiang (1992) and Černá and Chytrý (2005), artificial neural network (ANN) approach was selected to justify the result of statistical approaches.

3. Statistical Model Fitting Approach

3.1 Theory and Methodology

3.1.1 Multivariate Generalized Linear Models

In analogy to the univariate case, multivariate generalized linear models are based on a distributional assumption and a structural assumption (Fahrmeir and Tutz, 1994). However, the response variable y_i is now a q-dimensional vector with expectation $\mu_i = E(y_i|x_i)$.

1. Distributional assumption:

Given x_i, the y_i's are (conditionally) independent and have a distribution that belongs to a simple exponential family, which has the form with ω_i nuisance parameters and ϕ additional dispersion parameter.

$$f(y_i|\theta_i, \phi, \omega_i) = exp\left\{\frac{[y_i'\theta_i - b(\theta_i)]}{\phi}\omega_i + c(y_i, \phi, \omega_i)\right\} \tag{3.1}$$

2. Structural assumption:

The expectation μ_i is determined by a linear predictor

$$\eta_i = Z_i\beta \tag{3.2}$$

in the form

$$\mu_i = h(\eta_i) = h(Z_i\beta) \tag{3.3}$$

where

- The response function $h : S \rightarrow M$ is defined as $S \subset \mathbb{R}^q$, taking values in the admissible set $M \subset \mathbb{R}^q$,
- Z_i is a $(q \times p)$-design matrix, and
- $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)$ is a vector of unknown parameters from the admissible set $\boldsymbol{\beta} \subset \mathbb{R}^p$.

For the case of a multicategorical response, one has to consider the multinomial distribution, which may be embedded into the framework of a simple (multivariate) exponential family. For $y_i = (y_{i1}, y_{i2}, \dots, y_{iq}) \sim M(n_i, \pi_i)$ the distribution of the arithmetic mean $\bar{y}_i = {y_i}/{n_i}$ has the form

$$f(\bar{y}_i|\theta_i, \phi, \omega_i) = exp\left\{\frac{[\bar{y}_i'\theta_i - b(\theta_i)]}{\phi}\omega_i + c(y_i, \phi, \omega_i)\right\} \tag{3.4}$$

where the natural parameter θ_i is given by

$$\theta_i' = \left[log\left(\frac{\pi_{i1}}{1-\pi_{i1}-\pi_{i2}-\dots-\pi_{iq}}\right), \dots, log\left(\frac{\pi_{iq}}{1-\pi_{i1}-\pi_{i2}-\dots-\pi_{iq}}\right)\right] \tag{3.5}$$

and

$$b(\theta_i) = log(1 - \pi_{i1} - \pi_{i2} - \dots - \pi_{iq}) \tag{3.6}$$

$$c(y_i, \phi, \omega_i) = log\left(\frac{n_i!}{y_{i1}! \dots y_{iq}!(n - y_{i1} - y_{i2} - \dots - y_{iq})}\right) \tag{3.7}$$

$$\omega_i = n_i \tag{3.8}$$

The parameter ϕ may be treated as an additional dispersion parameter, and it is considered fixed with $\phi = 1$.

For the multinomial distribution, the conditional expectation μ_i is the vector of probabilities$\pi(x_i) = \left(\pi_{i1}, \pi_{i2}, \dots, \pi_{iq}\right)$, $\sum_{r=1}^{q} \pi_{ir} < 1$.Therefore the admissible set of expectations M is given by

$$M = \left\{\left(z_1, z_2, \dots, z_q\right) \middle| 0 < z_i < 1, \sum_i z_i < 1\right\}. \quad (3.9)$$

3.1.1.1 Polytomous Response Variables

The response variable may consist of more than two categories. The model under binary response can be extended to handle both Polytomous and Multiple Response Variables.

Suppose there is a single *I* categorical response variable and two explanatory variables with *J* and *K* categories respectively. To describe the relationship between the response and explanatory variables, a set of *(I-1)*logit models is required.

Let, P_{ijk}= Pr [response *i* for an observation with explanatory variables *(j,k)*]

$$Q_{ijk} = \sum_{m=1}^{i} P_{mjk}. \quad (3.10)$$

Then,

$$R_{ijk} = \frac{P_{ijk}}{1-Q_{ijk}} = \frac{P_{ijk}}{\sum_{m>i}^{I} P_{mjk}} \quad (3.11)$$

(3.11) is said to be the i^th^ continuation ratio for group (j, k). The (I-1) logits are defined as $log_e(R_{ijk})$; $i = 1, 2, \dots, (I-1)$; $j = 1, 2, \dots J$; $k = 1, 2, \dots K$.

Sometimes the response variable may have an order. In that case response variable should be considered as an ordinal variable. There are 3 logit models available for the ordinal responses.

1. Proportional Odds Model

$$log\left(\frac{\pi_1 + \cdots + \pi_j}{\pi_{j+1} + \cdots + \pi_J}\right) = \boldsymbol{X\beta_j}, \quad \text{j = 1,…,J-1.}$$

where π_j=probability of jth level of response variable, $\boldsymbol{\beta_j} = (\beta_{0j}\beta_{1j} \dots \beta_{mj})^T$, m= number of covariates, $\boldsymbol{X}$ = Model matrix.

Eg: For J=4

1. $log\left(\frac{\pi_1}{\pi_2+\pi_3+\pi_4}\right) = \boldsymbol{X\beta_1}$
2. $log\left(\frac{\pi_1+\pi_2}{\pi_3+\pi_4}\right) = \boldsymbol{X\beta_2}$
3. $log\left(\frac{\pi_1+\pi_2+\pi_3}{\pi_4}\right) = \boldsymbol{X\beta_3}$

2. Continuation Ratio Model

$$log\left(\frac{\pi_1+\cdots+\pi_j}{\pi_{J+1}}\right)=\boldsymbol{X\beta_j},\quad j = 1,\ldots,J\text{-}1.$$

where π_j=probability of jth level of response variable, $\boldsymbol{\beta_j}=(\beta_{0j}\beta_{1j}\ldots\beta_{mj})^T$, m= number of covariates, $\boldsymbol{X}$ = Model matrix.

Eg: For J=4

1. $log\left(\frac{\pi_1}{\pi_2}\right)=\boldsymbol{X\beta_1}$
2. $log\left(\frac{\pi_1+\pi_2}{\pi_3}\right)=\boldsymbol{X\beta_2}$
3. $log\left(\frac{\pi_1+\pi_2+\pi_3}{\pi_4}\right)=\boldsymbol{X\beta_3}$

3. Adjacent Categories Model

$$log\left(\frac{\pi_{j+1}}{\pi_j}\right)=\boldsymbol{X\beta_j},\quad j = 1,\ldots,J\text{-}1.$$

where π_j=probability of jth level of response variable, $\boldsymbol{\beta_j}=(\beta_{0j}\beta_{1j}\ldots\beta_{mj})^T$, m= number of covariates, $\boldsymbol{X}$ = Model matrix.

Eg: For J=4

1. $log\left(\frac{\pi_1}{\pi_2}\right)=\boldsymbol{X\beta_1}$
2. $log\left(\frac{\pi_2}{\pi_3}\right)=\boldsymbol{X\beta_2}$
3. $log\left(\frac{\pi_3}{\pi_4}\right)=\boldsymbol{X\beta_3}$

3.1.1.2 Adjacent Categories Model

Adjacent categories logit model is an alternative model which is applicable when the response variable is in ordinal scale, having more than two categories, and it compares a class with its adjacent class. Adjacent categories logits can be defined as $log\left[\frac{P_i}{P_{i+1}}\right]$. If there are *I* categories of the response variable, then there exists *(I-1)* number of adjacent categories logits. For each of these, *(I-1)* different regression models can be fitted.

$$log\left[\frac{P_{ijk}}{P_{(i+1)jk}}\right]=\alpha_i+\beta_{ij}^A+\beta_{ik}^B\ . \qquad (3.12)$$

For each model, parameters can be estimated separately. Each model can be

ISBN-1391-4987

considered as a binary logistic model when for each model $P_{ijk} + P_{(i+1)jk} = 1$ constraint has been used.

3.1.1.3 Maximum Likelihood Estimation

In the following, estimation is described in the general form of multivariate exponential families. The special case of the multinomial distribution is given by $f(\bar{y}_i|\theta_i, \phi, \omega_i) = exp\left\{\frac{[\bar{y}_i'\theta_i - b(\theta_i)]}{\phi}\omega_i + c(y_i, \phi, \omega_i)\right\}$. Based on exponential family

$$f(y_i|\theta_i, \phi, \omega_i) = exp\left\{\frac{[y_i'\theta_i - b(\theta_i)]}{\phi}\omega_i + c(y_i, \phi, \omega_i)\right\}. \quad (3.13)$$

For observation vector $y_i, \ldots, y_n$, maximum likelihood estimation may be derived in analogy to the one-dimensional case. The log-likelihood kernel for observation y_i is given by

$$l_i(\mu_i) = \frac{y_i'\theta_i - b(\theta_i)}{\phi}\omega_i, \quad \theta_i = \theta_i(\mu_i),$$

and the log likelihood for the sample has the form

$$l(\beta) = \sum_{i=1}^{n} l_i(\mu_i). \quad (3.14)$$

Note that in the grouped case n stands for the number of groups, whereas in the ungrouped case n stands the number of units. Using the link $\mu_i = h(Z_i\beta)$, the score function $s(\beta) = \partial l/\partial\beta = \sum_{i=1}^{n} s_i(\beta)$ has components

$$s_i(\beta) = Z_i' \; D_i(\beta)\,\Sigma_i^{-1}(\beta)[y_i - \mu_i(\beta)] \quad (3.15)$$

where

$$D_i(\beta) = \frac{\partial h(\eta_i)}{\partial\eta} \quad (3.16)$$

is the derivative of $h(\eta)$ evaluated at $\eta_i = Z_i\beta$ and

$$\Sigma_i(\beta) = cov(y_i) \quad (3.17)$$

denotes the covariance matrix of observation y_i given parameter vector β. The alternative form

$$s_i(\beta) = Z_i' \; W_i(\beta)\frac{\partial g(\mu_i)}{\partial\mu'}[y_i - \mu_i(\beta)] \quad (3.18)$$

makes use of the weight matrix

$$W_i(\beta) = D_i(\beta)\,\Sigma_i^{-1}(\beta)D_i'(\beta) = \left\{\frac{\partial g(\mu_i)}{\partial\mu'}\Sigma_i(\beta)\frac{\partial g(\mu_i)}{\partial\mu}\right\}^{-1} \quad (3.19)$$

which may be considered as an approximation of the inverse of the covariance matrix of the "transformed" observation $g(y_i)$ in the case where $g(y_i)$ exists. The expected Fisher information is given by

$$F(\beta) = cov\big(s(\beta)\big) = \sum_{i=1}^{n} Z_i' \; W_i(\beta)\, Z_i. \tag{3.20}$$

In matrix notation, score function and Fisher matrix are

$$s(\beta) = Z' D(\beta) \Sigma^{-1}(\beta)[y - \mu(\beta)], \qquad F(\beta) = Z'\, W(\beta)\, Z$$

respectively, where y and $\mu(\beta)$ are given by

$$y' = (y_1', \ldots \ldots, y_n'), \quad and \quad \mu(\beta)' = (\mu_1(\beta)', \ldots \ldots, \mu_n(\beta)').$$

The matrices have block diagonal form

$$\Sigma(\beta) = diag\big(\Sigma_i(\beta)\big), \quad \mathrm{W}(\beta) = diag\big(\mathrm{W}_i(\beta)\big), \quad \mathrm{D}(\beta) = diag\big(\mathrm{D}_i(\beta)\big)$$

and the total design matrix is given by $Z = [Z_1, Z_2, \ldots, Z_n]'$

These formulas are given for individual observations $y_1, \ldots, y_n$. For grouped observations, which are more convenient for computational purposes, the formulas are the same. The only difference is that the summation is over the grouped observations $y_1, \ldots, y_g$, where y_i is the mean of n_i observations, and $\Sigma_i(\beta)$ is replaced by $\Sigma_i(\beta)/n_i$.

Under regularity assumptions (comparable to the assumptions univariate generalized linear model) one gets asymptotic normality of the estimate

$$\hat{\beta} \overset{a}{\sim} N\left(\beta, F^{-1}(\hat{\beta})\right)$$

That means $\hat{\beta}$ is approximately normal with covariance matrix $cov(\hat{\beta}) =$ $F{-}1\beta$.

3.1.1.4 Model Selection Procedure

Forward selection method was used to select the best model among a set of hierarchical models. This procedure starts with the null model (intercept term only) and main effects are added one at a time and select the most significant factor (i.e. the term which gives minimum p value) to the model. Then, add the rest of the main effects to the previously selected model and select the next most significant term. This process continues until none of the main affects show significant improvement when it is adjusted for already selected variables. After selecting the model, then diagnostics of the model should be carried out to examine the adequacy of the fitted model.

3.1.1.5 Comparing Alternative Logistic Models

Alternative logistic models can be compared by examining the difference in deviance. When one model contains terms that are additional to the terms of another, the difference in the deviance measures the extent to which the additional term improves the fit of the model. The hypothesis that should be tested is,

- H_0: Additional term(s) does/do not improve the fit of the model significantly
- H_1: Additional term(s) improve(s) the fit of the model significantly

Let the deviance of the model 1 to be D_1 with p degrees of freedom and the deviance of model 2 to be D_2 with k degrees of freedom, where $p < k$. If reduction in deviance ($D = D_1$-D_2) is significant at α% level (i.e.D $> \chi^2_{(k-p),\alpha\%,}$), then model 2 is preferred over the model 1. That implies the additional terms should be in the model.

3.1.2 Testing a Goodness-of-fit

When testing of linear hypotheses of the form $H_0 : C\beta = \xi$ against $H_1 : C\beta \neq \xi$ one only has to substitute score functions and Fisher matrices by their multivariate versions.

3.1.2.1 Goodness-of-fit Statistics

Goodness-of-fit of the models may again be checked by the Pearson statistic and the deviance. Since we are considering multinomial data the expectation μ_i is equivalent to the probability vector $\pi_i = (\pi_{i1}, \ldots, \pi_{iq}), \quad \pi_{ik} = 1 - \pi_{i1} - \cdots - \pi_{iq}$, where $\pi_{ir} = P(Y = r|x_i)$. The estimates of π_i based on the model denoted by $\hat{\mu}_i = \hat{\pi}_{ik} = (\hat{\pi}_{i1}, \ldots, \hat{\pi}_{iq})$.

The Pearson statistic in general is given by

$$\chi^2 = \sum_{i=1}^{g} (y_i - \hat{\mu}_i)' \Sigma_i^{-1} \left(\hat{\beta}\right) (y_i - \hat{\mu}_i). \tag{3.21}$$

In the case of a multi-categorical response variable with multinomial distribution $n_i y_i \sim M(n_i, \pi_i) \chi^2$ may be written in the more familiar form

$$\chi^2 = \sum_{i=1}^{g} \chi_P^2 (y_i, \hat{\pi}_i) \tag{3.22}$$

where

$$\chi_P^2 (y_i, \hat{\pi}_i) = n_i \sum_{j=1}^{k} \frac{\left(y_{ij} - \hat{\pi}_{ij}\right)^2}{\hat{\pi}_{ij}} \tag{3.23}$$

is the Pearson residual for the ith (grouped) observation with $y_{ik} = 1 - y_{i1} - \cdots - y_{iq}, \hat{\pi}_{ik} = 1 - \hat{\pi}_{i1} - \cdots - \hat{\pi}_{iq}$.

The deviance or likelihood ratio statistic is given by

$$D = -2 \sum_{i=1}^{g} \{l_i(\hat{\pi}_i) - l_i(y_i)\} \quad (3.24)$$

For multinomial data the more familiar form is given by

$$D = 2 \sum_{i=1}^{g} \chi_D^2 (y_i, \hat{\pi}_i) \quad (3.25)$$

where

$$\chi_D^2(y_i, \hat{\pi}_i) = n_i \sum_{j=1}^{k} y_{ij} \log \left(\frac{y_{ij}}{\hat{\pi}_{ij}}\right) \quad (3.26)$$

is the deviance residual. If $y_{ij} = 0$ the term $y_{ij} \log(y_{ij}/\hat{\pi}_{ij})$ is set to zero.

Under "regularity conditions" including in particular increasing sample sizes $n_i \to \infty, i = 1, \dots, g$ such that $n_i/n \to \lambda_i > 0$ one gets approximately χ^2 -distributed goodness-of-fit statistics

$$\chi^2, \mathrm{D} \overset{a}{\sim} \chi^2(g(k-1) - p)$$

where$g$ denotes the number of groups, k is the number of response categories and p is the number of estimated parameters.

3.1.2.2 Goodness of Parameters

In particular p-values corresponding to the squared t-values t_r^2 of effects β_r, are computed from the $\chi^2(1)$ distribution

where

$$Wald\ Statistic = t_r^2 = \frac{\hat{\beta}_r}{\sqrt{var(\hat{\beta}_r)}} \sim \chi^2(1) \quad (3.27)$$

Definition of Residuals

A residual provides a measure of the extent to which each observation deviates from its fitted value under the assumed model. Depending on the specifications, there exist several types of residuals. The Pearson residual is the raw residual divided by the square root of the variance function. By standardizing Pearson's residuals, Standardized Pearson's residuals can be obtained. This is given by,

$$r_{P_i} = \frac{y_i - n_i\hat{p}_i}{\sqrt{n_i\hat{p}_i(1-\hat{p}_i)(1-h_i)}}. \quad (3.28)$$

4. Statistical Model Fitting

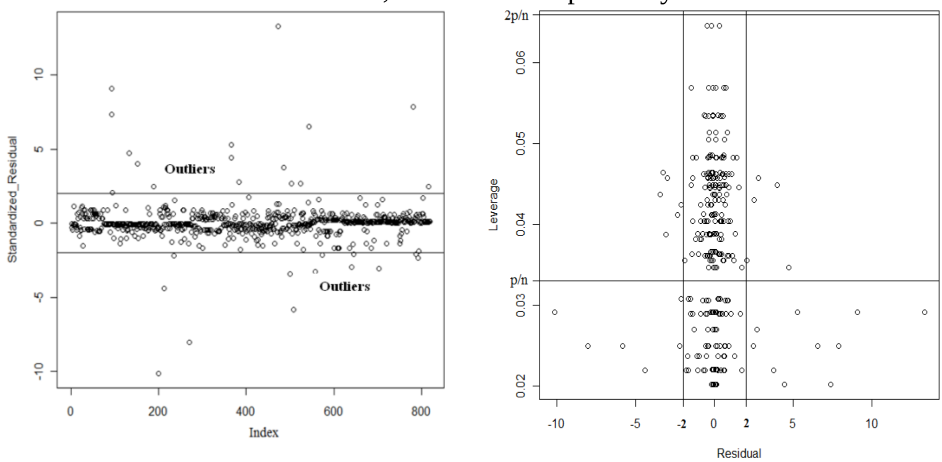
The response variable has four categories; senior management (M1), middle management (M2), junior executive (M3) and non executive grade (M4). ACL model compares the adjacent categories of ordered dependent variable. Hence, a multivariate generalized linear model can be fitted. Here, the full preprocessed data set is used for the modeling procedure.

The model for the data has been obtained by using *acat* function in a *VGLM* library (Yee, 2010) in the R statistical software package. It provides deviance statistics and t-values for the parameter estimates. In our study, the deviance statistics of the fitted ACL model is significantly very low, and the fitted model has very less number of outliers and influential values. Therefore it has enough evidence for the adequacy of the model.

Final ACL Model:

$$\begin{cases} \log\left(\frac{P(Y=M1)}{P(Y=M2)}\right) = \alpha_1 + \beta_{11}^{X8} + \cdots + \beta_{1J}^{X8} + \beta_{11}^{X5} + \cdots + \beta_{1K}^{X5} + \beta_{11}^{X2} + \cdots + \beta_{1L}^{X2} & (3.29) \\ \log\left(\frac{P(Y=M2)}{P(Y=M3)}\right) = \alpha_2 + \beta_{21}^{X8} + \cdots + \beta_{2J}^{X8} + \beta_{21}^{X5} + \cdots + \beta_{2K}^{X5} + \beta_{21}^{X2} + \cdots + \beta_{2L}^{X2} & (3.30) \\ \log\left(\frac{P(Y=M3)}{P(Y=M4)}\right) = \alpha_3 + \beta_{31}^{X8} + \cdots + \beta_{3J}^{X8} + \beta_{31}^{X5} + \cdots + \beta_{3K}^{X5} + \beta_{31}^{X2} + \cdots + \beta_{3L}^{X2} & (3.31) \end{cases}$$

where β_{ij}^{Xk} is the j^{th} level effect over the base level of the k^{th} covariate for the i^{th} logit (dummy coded variable) and α_i is the intercept of the i^{th} equation. J, K and L are the number of levels in X8, X5 and X2 respectively.



(a) (b)

Figure 1: (a) Standardized residuals of ACL Model(b) Leverage Values of ACL Model

The above figure 1(a) shows the standardized residual plot of the fitted ACL model, and these residuals are calculated for each logits in the model. Most of the residuals lied between (+2,-2) region and considerably very small number of residuals have the value beyond (+2,-2). The residuals which have the value beyond (+2,-2) are considered as outliers and usually a large number of outliers indicates that the fitted model is inadequate, However the number of outliers is considerably a very small number in this study, which is an evidence to the adequateness of the model. Most of the residuals lied around the zero value.

The final fitted model's deviance is 361.209312 with 792 degrees of freedom. Usually a higher value of deviance indicates inadequacy of the model. However, in this study, it is significantly very low with respect to the degrees of freedom. Therefore it can be concluded that the fitted model fits the data well.

Leverage measures the potential impact of an individual case on the results, which is directly proportional to how far an individual case is from the centre in the space of the predictors. Leverage is computed as the diagonal elements of the hat matrix. To detect the high leverage points, leverage values can be plotted. The cutoff value for leverage can be taken as twice or thrice as large as the average leverage. The average leverage value is given by p/n where n is the number of observations and p is the number of estimated parameters. The variables present in the ACL model are X8 with four levels, X5 with four levels and X2 with three levels. There are three parallel logits, therefore p =3*(1+ (4-1) + (4-1) + (3-1)) = 27 and n = 3*273 gives the three cut off values p/n=0.033, $2p/n$=0.066 and $3p/n$=0.099. According to Figure 5(b), we can observe that there exist some points which lie beyond the lower cutoff value p/n=0.03296703 but none of them exceeds the upper critical value $2p/n$=0.06593407. Therefore it is reasonable to conclude that there are no extreme points in the data.

Table 2 given below shows the summary of the final fitted adjacent category logit model. The model includes three parallel logits, as three logit models with only three significant explanatory variables such as sector of the first job (X5), organization of second job (X8) and the university (X2), among 16 variables which are considered in this study. The table clearly shows that the variable organization of second job(X8) has a significant influence to the entire model since many levels of that variable in all the sub models are significant. However in other two variables, at least one level of each variable among three sub models has a significant effect.

Table 1: Parameter Estimates of ACL Model

Variable	Level of Variable	Model No	Value	Std. Error	t value	P-Value	
Intercept		1	3.567186	1.23866	2.87988	0.007957	**
		2	-0.325140	1.23894	-0.26244	0.414036	
		3	-4.224360	0.93982	-4.49484	0.000014	***
X5 (Sector of First Job)	Government Sector	1	0	0	0	0	
		2	0	0	0	0	
		3	0	0	0	0	
	Semi Government Sector	1	0.771541	0.99839	0.77279	0.879293	
		2	0.475955	0.91588	0.51967	0.793413	
		3	0.792377	0.78379	1.01095	0.624081	
	Private Sector	1	-0.657580	1.00753	-0.65267	0.972062	
		2	-1.315090	0.85972	-1.52968	0.252192	
		3	3.557222	0.76202	4.66813	0.000006	***
	Other Sector	1	-2.021490	1.09583	-1.84471	0.130159	
		2	1.770630	0.94063	1.88239	0.119566	
		3	0.864120	0.91285	0.94662	0.687665	
X8 (Organization of Second Job)	Academic Organization	1	0	0	0	0	
		2	0	0	0	0	
		3	0	0	0	0	
	Consultancy Organization	1	-2.221850	1.42083	-1.56377	0.235743	
		2	-3.736250	1.38302	-2.70151	0.013805	*
		3	5.278538	0.82388	6.40696	0	***
	Trade & Business Organizations	1	-3.732960	0.98797	-3.77841	0.000316	***
		2	-0.668170	1.01543	-0.65802	0.978949	
		3	5.029467	0.94981	5.29524	0	***
	Other Organizations	1	-6.095990	1.10091	-5.53722	0	***
		2	0.859881	0.99966	0.86018	0.779380	
		3	5.840110	1.22403	4.77122	0.000004	***
X2 (University)	Central Universities	1	0	0	0	0	
		2	0	0	0	0	
		3	0	0	0	0	
	Western Universities	1	0.190879	0.71833	0.26573	0.419106	
		2	0.206657	0.60561	0.34124	0.534154	
		3	0.093971	0.67241	0.13975	0.222285	
	Other Universities	1	3.032586	1.26339	2.40036	0.032758	*
		2	-3.297040	1.16498	-2.83012	0.009306	**
		3	0.595501	0.73409	0.81121	0.834490	

*Note: Significant Levels: “***”⟹0.1%, “**”⟹1%, “*”⟹5%*

Findings reveal that;

1. A person who is employed in an academic organization has about five hundred times higher chance of obtaining a senior managerial position than middle managerial position relative to a person who employed by another organization, as second job position.
2. A person who has graduated in a university which belongs to the "other universities"category has just about twenty times higher chance of obtaining a senior managerial position than middle managerial position relative to a person who has graduated from central university, as second job position.
3. A person who is employed in an academic organization has just about fifty times higher chance of obtaining a middle managerial position than junior managerial position relative to a person who employed in a consultancy organization, as second job position.
4. A person who has graduated in central university has just about thirty times higher chance of obtaining a middle managerial position than junior managerial position relative to a person who has graduated in a university which belongs to the "other universities" category as second job position.
5. A person who did his first job in private sector has about thirty five times higher chance of obtaining a senior managerial position than non executive grade position in his second job than that for a person who did his first job in government sector, as second job position.
6. A person who is employed in a consultancy organization has just about two hundred times higher chance of obtaining a junior managerial position than non executive grade position relative to a person who employed in an academic organization, as second job position.
7. A person who would be employed in an organization which belongs to the "other organizations" category has just about three hundred and fifty times higher chance of obtaining a junior managerial position than non executive grade position relative to a person who employed in an academic organization, as second job position.
8. A person who would be employed in an organization which belongs to the "trade and business organizations" category has just about hundred and fifty times higher chance of obtaining a junior managerial position than non executive grade position relative to a person who employed in an academic organization, as second job position.

Here some contradicting results can be noticed in a practical situation, which may be due to insufficient data for some particular categories in the variable. If we have more data then we may have conclusions which make more sense.

5. Neural Network Approach

A best fitted neural network can be used for predictions. Among 291 row data 212 cases were assigned to the training sample, and 79 to the testing sample. The constructed best neural network model has one input layer with 15 units (4 factors), one hidden layer with 6 units and one output layer with 4 units (one dependent variable). The input layer has organization of second job (X8), sector of the first job (X5), sector of the second job (X9) and the field name of the second job (X10).The output layer has a second job position(Y) as factors. The *hyperbolic tangent* is the activation function which is used in the hidden layer, and the *softmax* activation function is used in output layer.

The accuracy of the classification of the fitted ANN model is as follows. The training data has a 77.8 percentage of accuracy and the testing data has a 77.2 percentage of accuracy. Those numbers indicate that the neural network is significantly trained and the fitted model can classify more than 3/4th of observations accurately. The corresponding ROC curves for all four categories are symmetrical about a 45-degree line from the upper left corner of the chart to the lower right. They closely follow the left-hand border, and then the top border of the ROC space, which indicate the higher accuracy.

The importance of an independent variable is a measure of how much the network's model-predicted value changes for different values of the independent variable. Normalized importance values are simply the importance values divided by the largest importance value, and expressed them as percentages. The below importance chart is simply a bar chart of the values in the importance table, sorted in descending value of importance. In Figure 2, it appears that variables X8 and X5 have the greatest effect on how the network classifies second job position.

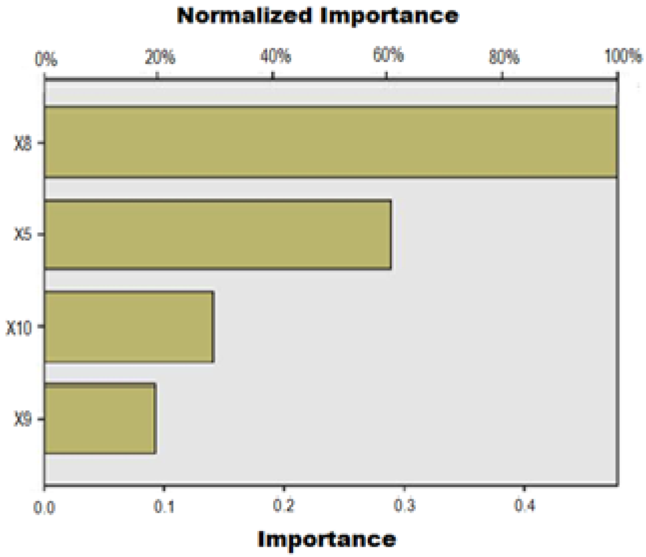


Figure 2: Independent Variable Importance

6. Conclusion

This study reveals that “organization of second job (X8)”, “sector of the first job (X5)” and “University (X2)” are the factors that associate with the second job position. Results from the statistically fitted model have been justified using ANN approach. ACL model, as a rarely fitted model, can provide useful and more intuitive results that can be used to improve an individual’s future job positions.

Specially, people who are working in human resources department can benefit from this study. This study can be done not only for the second job position, but also for the general case such as next job position. One can consider not only the managerial positions, but also for any other job criteria. Relevant tons of data can be collected from professional social networks such as www.linkedin.com. It is very useful for every company to overcome problems in human resource management if they have a model like this.

Not only in this context, there are many other real world problems that can be analyzed using ACL models. This study is a good example to motivate people to use ACL model rather than paying attention only on classical approaches.

References

1. Agresti A. (1992). Analysis of Ordinal Paired Comparison Data. *Journal of the Royal Statistical Society, 41*(2):287-297.
2. Agresti A. (1996). Categorical Data Analysis (2nd ed.). New York: John Wiley & Sons.
3. Černá L. and Chytrý M. (2005). Supervised Classification of Plant Communities with Artificial Neural. *Journal of Vegetation Science, 16*(4):363-372, DOI:10.1111/j.1654-1103.2005.tb02380.x
4. Fahrmeir L. and Tutz G. (1994). Multivariate Statistical Modelling Based on Generalized Linear Models. New York: Springer-Verlag, DOI:10.1007/978-1-4899-0010-4
5. Ganzach Y. (1998). Intelligence and Job Satisfaction. *The Academy of Management Journal, 41*(5):526-539, DOI:10.2307/256940
6. Goodman L. A. (1983). The Analysis of Dependence in Cross-Classifications Having Ordered Categories Using Log-Linear Models for Frequencies and Log-Linear Models for Odds. *International Biometric Society, 39*(1):149-160, DOI:10.2307/2530815
7. Koh H. C. and Boo E. H. Y. (2001). The Link between Organizational Ethics and Job Satisfaction: A Study of Managers in Singapore. *Journal of Business Ethics, 29*(4):309-324, DOI:10.1023/A:1010741519818
8. Tam K. Y. and Kiang Y. M. (1992). Managerial Applications of Neural Networks: The Case of Bank Failure Predictions. *Management Science, 38*(7): 926-947, DOI: 10.1287/mnsc.38.7.926
9. Volkwein J. F. and Zhou Y. (2003). Testing a Model of Administrative Job Satisfaction. *Research in Higher Education, 44*(2):149-171, DOI:10.1023/A:1022099612036
10. Wharton A. S., Rotolo T. and Bird S. R. (2000). Social Context at Work: A Multilevel Analysis of Job Satisfaction. *Sociological Forum, 15*(1):65-90, DOI:10.1023/A:1007546120560
11. Yee T. W. (2010). The VGAM Package for Categorical Data Analysis. *Journal of Statistical Software, 32*(10).