

Relative importance of explanatory variables in a logistic regression

CHAO FU

04 January, 2022

Abstract

The relative importance of explanatory variables in multiple linear regression has received great attention for over half a century. A considerable amount of researches have been spent on researching and developing a unique, rational and interpretable relative importance indices. With the increasing popularity of logistic regression in more and more fields such as medicine, epidemic, psychology and economics, extensive attention has been prone to the relative importance of logistic regression. However, there are few investigations of interest to the influence on relative importance by interaction explanatory variables. This paper aims to compare the discrepancies between the model with and without interaction based on four relative importance indices: Pratt, Dominance analysis, Relative weights and Random forest.

Keywords: Logistic Regression, Relative Importance, Interactional Explanatory Variables

1 Introduction

Multiple linear regression is used to analyze a model with multiple independent variables and a dependent variable which all are continuous. In medicine, psychology and epidemic, analyzing the model with binary dependent variables has become prevalent in the last 20 years. Logistic regression is an appropriate method for this situation instead of multiple regression. The classic logistic regression model estimates the probability of each response variable value occurring based on a set of independent variables.[1]

Independent variables are also commonly referred to as explanatory variables, covariates, predictors, regressors and attributes. The dependent variable is also commonly referred to as response variable and response. Interaction explanatory variables mean that one predictor influences the relation between another predictor and response. Even though interaction explanatory variables can make the model more complete, they can raise concerns of multicollinearity in the model.

The basic multiple linear regression equation is as follows:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p = \beta^T X$$

The interpretation of this equation is given below:

- 1) \hat{Y} is the estimated continuous value matrix($1 \times N$).
- 2) β_0 is the intercept which is regarded as a constant value. X_1, X_2, \dots, X_p are p number of explanatory variables. $\beta_1, \beta_2, \dots, \beta_p$ are their regression coefficients. β is a regression coefficient matrix($(p+1) \times 1$) and X is a sample matrix($(p+1) \times N$) in which N is the number of observations.

The conventional logistic regression equation is as follows:

$$P(Y_i) = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p}}$$

Using logit to transform the basic logistic regression into a form of multiple linear regression is given below:

$$\text{logit}(\hat{Y}) = \ln(p/1 - p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \beta_p X_p = \beta^T X$$

The components of this equation are as follows:

- 1) $p = P(Y_i)$ is the estimated probability of each response variable value occurring.
- 2) Others are the same as multiple linear regression noted before.

Multiple regression analysis has two different purposes: estimating the dependent variable value and exploring the effects of independent variables on the dependent one[2]. This also carries over to logistic regression analysis.

Achen debated three distinct variable importance given below[3].

Name	Meaning
Theoretical importance	Regression coefficients
Level importance	Product of a variable's mean and its unstandardized regression coefficients
Dispersion importance	Amount of explained variance

With respect to dispersion importance, R^2 , coefficient of determination(also called squared multiple correlation coefficient), is prevalent in relative importance due to quantifying predictability and giving the strength of a regression relationship. Many measures of explained variance for logistic regression based on R^2 has been developed[4].

Standardized coefficients should be applied in relative importance measurements based on R^2 . The key reason is that when variables are measured in different units of measurement, standardized coefficients are beneficial for the comparison of the relative importance among different predictors in both multiple linear and logistic regression models [5].

Standardized coefficients for multiple regression are simple, but hard for logistic regression. Menard reviewed four main alternative methods to calculate standardized logistic coefficients, given below[6]:

No.	Notion
1	$\beta_A^* = \beta_1 sd(X_1), \beta_2 sd(X_2), \dots \beta_p sd(X_p)$
2	$\beta_S^* = \beta_1 sd(X_1)/(\pi/\sqrt{3}), \beta_2 sd(X_2)/(\pi/\sqrt{3}), \dots \beta_p sd(X_p)/(\pi/\sqrt{3})$
3	$\beta_L^* = \beta_1 sd(X_1)/[(\pi/\sqrt{3}) + 1], \beta_2 sd(X_2)/[(\pi/\sqrt{3}) + 1], \dots \beta_p sd(X_p)/[(\pi/\sqrt{3}) + 1]$
4	$\beta_M^* = \beta_1 sd(X_1)R/sd(\text{logit}(\hat{Y})), \beta_2 sd(X_2)R/sd(\text{logit}(\hat{Y})), \dots \beta_p sd(X_p)R/sd(\text{logit}(\hat{Y}))$

sd means standard deviation.

Determining the relative importance of predictors is uncomplicated when they are independent of each other. However, it can become a surprisingly tough problem when there exists multicollinearity which can occur when two or more predictors are highly correlated in both multiple and logistic regression models [7]. An extreme example is given below:

$$\begin{aligned} y &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \\ X_1 &= bX_2 \\ y &= \beta_0 + (b\beta_1 + \beta_2)X_2 + \beta_3 X_3 \end{aligned}$$

In this case, coefficient $b\beta_1 + \beta_2$ can be estimated accurately as a whole. However, the estimation of β_1 and β_2 respectively can be unstable. Even though it does not influence the model's stability and ability, it can estimate coefficients volatile and provide a misleading result of relative importance among predictors.

There are two general methods to determine serious multicollinearity among predictors. One is to use a correlation matrix for explanatory variables. The large correlation coefficient for two predictors is regarded as multicollinearity. The other one is to use variance inflation factor(VIF)[8]. If $VIF = 1$, it suggests that predictors are independent of each other. Multicollinearity occurs when $VIF > 10$ [9] or > 5 [10].

Hence, it is no surprise that exploring a unique, rational and interpretable relative importance method has been targeted as a crucial task for over half a century. Fortunately, many state-of-art methods have been developed.

2 Background

There are two types for multiple linear regression to measure relative importance[11]:

- 1) Treat domains as independent variables
- 2) Treat domains as dependent variables, such as multivariate analysis of variance(MANOVA)

The second type is not included in this paper. For the first type, a large number of relative importance methods over the last 60 years has been developed[12]. With a significantly growing demand for logistic regression in sociology, economics, psychology, epidemic, medicine and ecology, the research of relative importance of logistic regression has exploded in development in the last 20 years. Four distinct methods are applied in this paper:

Name	Multiple regression	Year	Logistic regression	Year	package	Type
Pratt	Pratt	1987	Thomas et al.	2008	relaimpo	Single-Analysis
Dominance Analysis	Azen and Budescu	2003	Traxel and Azen	2009	dominanceanalysis	Multiple-Analysis
Relative Weights	Johnson	2000	Tonidandel and LeBreton	2010	-	Variable-Transform
Random Forest	Breiman	2001	Breiman	2001	randomForest	Machine Learning

“relaimpo”, R package, can not work on the logistic regression model directly. It can perform on the transformed logistic regression model by logit. Moreover, it can not work on the model with interaction explanatory variables. Hence, a new algorithm is designed based on the Pratt method to fix these problems, seen in the Appendix.

There is no R package for Tonidandel and LeBreton’s method. Hence, another new algorithm is designed, seen in the Appendix.

The description of different types is as follows[13]:

Method	Description
Single-Analysis	Use the output from a single regression analysis, either by choosing a single index to represent the importance of the predictors or by combining multiple indices to compute a measure of importance
Multiple-Analysis	Combine the results from more than one regression analysis involving different combinations of the same variables
Variable Transformation	Transform the original predictors to a set of uncorrelated variables and regress the criterion on the uncorrelated variables, either use those results as a proxy for inferring the importance of the original variables or further analyze those data to yield results that are directly tied to the original variables

1) Pratt

Pratt provided a theoretical and axiomatic measurement of the relative importance of explanatory variables. There are a set of assumptions and axioms that should be satisfied to yield a unique measure. The main part of the Pratt method for logistic regression model is given below:

$$\text{logit}(\hat{Y}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \beta_p X_p$$

$$R^2 = \sum_{i=1}^p b_i \rho_i$$

- (1) β_0 is intercept, $\beta_1, \beta_2, \dots \beta_p$ are unstandardized regression coefficients.
- (2) $b_i \rho_i$ is defined as the measurement of predictors X_i importance.
- (3) b_i is the standardized regression coefficients of X_i . It is obtained by the fourth method mentioned before in my algorithm.
- (4) ρ_i is the correlation coefficients between X_i and $\text{logit}(\hat{Y})$.

Pratt stated that both b_i and ρ_i must have the same sign meaning that the predictors' importance must be positive[14]. This statement led to a large number of criticisms. Thomas et al. declared that negative importance only exists in the multicollinearity case based on the Pratt method. But negativity is a sufficient but not necessary indicator of multicollinearity. They also proved that the Pratt method has a negative lower bound for the negative importance[15], as follows:

$$b_i \rho_i \geq -(1/2)[(VIF_i)^{1/2} - 1]R^2, i = 1, 2, \dots p$$

Hence, the negative importance is not uncommon in the Pratt method. But Hoerl and Kennard used ridge regression to obtain positive importance instead of the negative one in the Pratt method. They stated that the original negative importance is wrong[16].

Another criticism is that the Pratt method is non-intuitive. But Bring [17] and Thomas et al.[15] proved that Pratt's method is an intuitive definition.

2) Dominance Analysis

Azen and Budescu defined the dominance analysis method as a unique and intuitive way to determine the relative importance. All unordered pairs of predictors from the model are compared by all subset models excluding the paired ones. The change of R^2 is regarded as the additional contribution when each subset model is added to the compared predictors[18]. For example, a model contains 3 explanatory variables X_1, X_2, X_3 . The additional contribution for X_1 is $R^2_{Y.X_1Z} - R^2_{Y.X_1}(Z \in \{X_2, X_3, X_2X_3\})$

$R^2_{Y.X_1Z}$ means R^2 values of the model of $Y = X_1 + Z(Z \in \{X_2, X_3, X_2X_3\})$ respectively.

Razia and Traxel[19] extended dominance analysis to logistic regression. They reviewed four criteria for defining appropriate R^2 analogues in logistic regression model, given below:

Property	Description
Boundedness	The value should be in the range 0-1, 0 means worst fit and 1 means the best fit
Linear invariance	Non-singular linear transformation of variables should not impact the value
Monotonicity	The value never reduce with the addition of a predictor
Intuitive interpretability	The value is explained intuitively

and four R^2 analogs methods which satisfied at least three of these four properties given below:

Notion	Year	Author	Boundedness	Linear invariance	Monotonicity	Intuitive inter-pretability
$R_M^2 = 1 - \frac{\ln(L_M)}{\ln(L_0)}$	1974	McFadden	Yes	Yes	Yes	Yes
$R_N^2 = \frac{1 - (L_0/L_M)^{2/n}}{1 - (L_0)^{2/n}}$	1991	Nagelkerke	Yes	Yes	Yes	NO
$R_E^2 = 1 - \left[\frac{\ln(L_M)}{\ln(L_0)} \right]^{-(2/n)\ln(L_0)}$	1998	Estrella	Yes	Yes	Yes	Yes
$R_{(y,\hat{y})}^2 = 1 - \frac{\sum (y-\hat{y})^2}{\sum (y-\bar{y})^2}$	2000	Zheng, Agresti	Yes	Yes	No	Yes

There are three types of dominance analysis: complete dominance, conditional dominance and general dominance. The complete dominance should be measured first. If it can not be established, then the conditional dominance should be measured. If it can not work, general dominance should be applied finally.

3) Relative Weights

The core restriction for the assessment of relative importance is multicollinearity among predictors. The relative weights method was first developed for the multiple linear regression model by Johnson. This method is to find the same number of orthogonal and uncorrelated vectors from original predictors.

- (1) $X = P\Delta Q'$
- (2) $Z = PQ'$
- (3) $\beta^* = (Z'Z)^{-1}Z'y$
- (4) $\Lambda^* = (Z'Z)^{-1}Z'X$
- (5) $\epsilon_{OLS} = \Lambda^{*2}\beta^{*2}$

X is scaled sample matrix ($N \times P$) and is full rank. $X = P\Delta Q'$ is singular value decomposition of X . Z is the orthonormal approximation to X . β^* are the regression coefficients on Z which are independent of each other. Although β^{*2} can define the relative importance without uncertainty, they are only approximations of the relative importance of X . Λ^* are the coefficients between Z and X . ϵ is the real relative importance of X . [20]

Then Tonidandel and LeBreton extended this method to logistic regression model [21]:

- (1) $X = P\Delta Q'$
- (2) $Z = PQ'$
- (3) sdz = standard deviation of Z by column
- (4) $\Lambda^* = (Z'Z)^{-1}Z'X$
- (5) β^* unstandardized logistic regression coefficients based on Z .
- (6) $R_0^2 = 1 - \frac{\sum (y-\hat{y})^2}{\sum (y-\bar{y})^2}$, \hat{y} is the fitted value of logistic regression model.
- (7) $s_{logit(\hat{y})}$ = standard deviation of logit value. $logit(\hat{y}) = \ln[\hat{y}/(1-\hat{y})]$
- (8) b^* standardized logistic regression coefficients for Z by the fourth method as mentioned previously.
- (9) $\epsilon_{LR} = \Lambda^{*2}b^{*2}$

4) Random Forest

Random Forest, one of the cutting-edge machine learning algorithms, has an important feature for measurement of relative importance. It does not use the same method applied in the linear regression model but adopts two types of importance measures: one is mean decrease inaccuracy which is also called mean squared error (MSE) reduction. The other one is the mean decrease in node impurity which is also called the Gini value. To evaluate certain predictor importance, the values of this variable are permuted. Then, the mean decrease

accuracy or Gini represents its importance. The importance is increasing with the growth of these two types of values. The mean decrease Gini is applied in this paper. Compared with the linear regression model, the Random Forest algorithm has many merits for relative importance[22]: (1) nonparametric (2) containing interactions (3) unaffected by multicollinearity.

3 Numerical example

Worldwide, breast cancer is the most common type of cancer in women and the second highest in terms of mortality rates. Diagnosis of breast cancer is performed when an abnormal lump is found (from self-examination or x-ray) or a speck of calcium is seen (on an x-ray). After a suspicious lump is found, the doctor will conduct a diagnosis to determine whether it is cancerous. More importantly, the doctor wants to explore which aspects play an important role in causing breast cancer.

This *breast cancer dataset* was obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg proposed in Kaggle.

The dataset consists of 569 samples and 6 variables :

One depends on a variable, “diagnosis”, 0 means no breast cancer and 1 means binary breast cancer.

Five explanatory variables, mean value of “radius”, “texture”, “perimeter”, “area” and “smoothness” which are continuous.

Hence, logistic regression which is a kind of generalized linear regression can work on this problem instead of multiple linear regression.

Moreover, the dataset is balanced and doesn’t contain missing data.

4 Results

4.1 Relative importance without interactions between explanatory variables

Method	Radius	Texture	Perimeter	Area	Smoothness	R^2
VIF	370.9409	1.16596	365.4882	40.4182	1.5880	-
Pratt	-2.1174	0.0761	1.4373	1.4482	0.0718	$0.9159(R^2_{(y,\hat{y})})$
Dominance Analysis	0.1962(25%)	0.0834(11%)	0.2168(28%)	0.1890(24%)	0.0894(12%)	$0.7748(R^2_M)$
Relative weights	0.2182(25%)	0.0745(9%)	0.2219(26%)	0.2562(30%)	0.0813(10%)	$0.8521(R^2_{(y,\hat{y})})$
Random Forest(Gini)	55.9379	33.0170	82.4814	65.6018	28.8535	-

4.2 Relative importance with interactions between explanatory variables

Method	Radius	Texture	Perimeter	Area	Smoothness	perimeter:area	R^2
VIF	881.8528	1.1888	422.9487	1003.0387	1.5903	267.9924	-
Pratt	-5.6160	0.1068	2.7681	7.2436	0.1227	-3.7176	$0.9076(R^2_{(y,\hat{y})})$
Dominance Analysis	0.1566(20%)	0.0773(10%)	0.1743(22%)	0.1494(19%)	0.0894(11%)	0.1479(18%)	$0.7945(R^2_M)$
Relative weights	0.1628(19%)	0.1085(13%)	0.1813(21%)	0.1060(13%)	0.1218(14%)	0.1677(20%)	$0.8482(R^2_{(y,\hat{y})})$

Method	Radius	Texture	Perimeter	Area	Smoothness	perimeter:area R^2
Random For- est(Gini)	54.1038	31.8110	82.4932	68.1070	28.7904	- -

5 Discussion and Conclusion

- 1) “relaimpo”, R package, can not work on the logistic regression model directly. It can perform on the transformed logistic regression model by logit. Moreover, it can not work on the model with interaction explanatory variables. Hence, a new algorithm is designed based on the Pratt method to fix these problems.
- 2) There is no R package for Tonidandel and LeBreton’s method to measure the relative importance in logistic regression model. Hence, another new algorithm is designed, seen in the Appendix.
- 3) There is significant multicollinearity in the model without interaction for the large VIF value of predictors “radius”, “perimeter” and “area”. The interaction between “perimeter” and “area” makes VIF of “radius”, “perimeter” and “area” increase over 2, 1, 25 times respectively. This situation can display that interaction can lead to extensive multicollinearity in the model.
- 4) As a result of serious multicollinearity, the Pratt method obtains more negative importance. These results show that negative importance is not uncommon for the Pratt method.
- 5) Dominance Analysis, Relative Weights and Random Forest, as the state-of-the-art method, can overcome the Pratt limitation to measure all positive importance.
- 6) Pratt, Dominance Analysis and Relative Weights can score the interaction predictor’s importance except for Random Forest. With and without interaction, Random Forest can obtain almost the same result. Its algorithm contains the interaction among predictors. Hence, Random Forest can be superior to other methods. However, Random Forest works like a “black box” for its unexplained results[23]. With this motive, more researches can be done to the improvement of this algorithm to make it more interpretable.
- 7) Without interaction, Dominance Analysis, Relative Weights and Random Forest can measure the same large and small magnitude predictors’ importance. However, the order is different.
- 8) Dominance Analysis can be a good method for measurement predictors’ importance. However, it has a significant drawback that the computation can have a considerable increase with the growth of predictors in the model. Hence, more attention should be given to overcome this shortcoming. Fortunately, Rady et al. has improved.[24]
- 9) Neural networks can have a surprisingly powerful computational ability. Hence, it can be beneficial for Dominance Analysis.
- 10) Bayesian network can be developed to measure predictors’ importance.

6 Appendix

```
knitr::opts_chunk$set(echo = TRUE, results = "markup")
rm(list=ls())
library(tidyverse)
library(ggplot2)
library(relaimpo)
library(dominanceanalysis)
library(randomForest)
```

```

# Load data
my_data <- read_csv("Breast_cancer_data.csv", col_names = TRUE, show_col_types = FALSE) %>% set_names("diagnosis")
# Pratt of relative importance indice without interaction
my_logre <- glm(diagnosis ~ ., family = binomial, data = my_data)
prob <- my_logre$fitted.values
new_data <- my_data %>% mutate(diagnosis = log(prob / (1 - prob)))
my_pratt <- calc.relimp(diagnosis ~ ., data = new_data, type = "pratt")
# Dominance analysis of relative importance indice without interaction
my_da <- dominanceAnalysis(my_logre)
r <- my_da$contribution.average$r2.m
# Relative weight of relative importance indice without interaction
my_rw <- function(r){
  P <- dim(r)[2]
  x <- scale(as.matrix(r[, -P]))
  y <- r$diagnosis
  de_svd <- svd(x)
  p <- de_svd$u
  qt <- t(de_svd$v)
  z <- scale(p %*% qt)
  v_star <- solve(t(z) %*% z) %*% t(z) %*% x
  new_r <- cbind(z, r[, P])
  colnames(new_r) <- colnames(r)
  new_x <- as_tibble(new_r)
  rw_logre <- glm(diagnosis ~ ., family = binomial, data = new_x)
  prob <- rw_logre$fitted.values
  b <- unname(rw_logre$coefficients[-1])
  sd_logit <- sd(log(prob / (1 - prob)))
  r0 <- sqrt(1 - (sum((y - prob) ^ 2) / sum((y - mean(y)) ^ 2)))
  beta_star <- b * r0 / sd_logit
  epsilon <- t(v_star ^ 2 %*% beta_star ^ 2)
  colnames(epsilon) <- colnames(r)[-P]
  return(epsilon)
}
# Random Forest of relative importance indice without interaction
new_data_rf <- my_data %>% mutate(diagnosis = as.factor(diagnosis))
my_rf <- randomForest(diagnosis ~ ., data = new_data_rf, importance = TRUE)
rf_vi <- my_rf$importance %>% as_tibble() %>% bind_cols(as_tibble(colnames(new_data)[-6])) %>% arrange(desc(importance))
# Pratt relative importance indice with interaction
my_logre1 <- glm(diagnosis ~ . + perimeter : area, family = binomial, data = my_data)
prob1 <- my_logre1$fitted.values
beta <- my_logre1$coefficients[-1]
new_data1 <- my_data %>% mutate(diagnosis = log(prob1 / (1 - prob1)))
x = as.matrix(new_data1[, -6])
perimeter_area <- x[, 3] * x[, 4]
new_x = cbind(x, perimeter_area)
y = as.matrix(new_data1[, 6])
y0 <- as.matrix(my_data[, 6])
new_xy <- cbind(new_x, y)
b <- cor(new_xy)[7, 1:6]
sd_logit <- sd(log(prob1 / (1 - prob1)))
sd_x <- apply(new_x, 2, sd)
r0 <- sqrt(1 - (sum((y0 - prob1) ^ 2) / sum((y0 - mean(y0)) ^ 2)))
beta_star <- beta * sd_x * r0 / sd_logit

```



```

my_pratt1 <- b * beta_star
vif_X <- as.matrix(new_x)
vif <- diag(solve(cor(vif_X)))
# Dominance analysis of relative importance indice with interaction
my_da1 <- dominanceAnalysis(my_logre1)
my_da1_r <- my_da1$contribution.average$r2.m
# Relative weight of relative importance indice with interaction
my_rw_1 <- function(r){

  P <- dim(r)[2]
  x1 <- scale(as.matrix(r[, -P]))
  x <- cbind(x1, x1[, 3] * x1[, 4])
  colnames(x) <- c(colnames(x1), "perimeter:area")
  y <- r$diagnosis
  de_svd <- svd(x)
  p <- de_svd$u
  qt <- t(de_svd$v)
  z <- scale(p %*% qt)
  v_star <- solve(t(z) %*% z) %*% t(z) %*% x
  new_r <- cbind(z, r[, P])
  colnames(new_r) <- c(colnames(x), "diagnosis")
  new_x <- as_tibble(new_r)
  rw_logre <- glm(diagnosis ~ ., family = binomial, data = new_x)
  prob <- rw_logre$fitted.values
  b <- unname(rw_logre$coefficients[-1])
  sd_logit <- sd(log(prob / (1 - prob)))
  r0 <- sqrt(1 - (sum((y - prob) ^ 2) / sum((y - mean(y)) ^ 2)))
  beta_star <- b * r0 / sd_logit
  epsilon <- t(v_star ^ 2 %*% beta_star ^ 2)
  colnames(epsilon) <- colnames(new_x)[-P-1]
  return(epsilon)
}
# Random Forest of relative importance indice with interaction
my_rf1 <- randomForest(diagnosis ~ . + perimeter : area, data = new_data_rf, importance = TRUE)
rf_vi1 <- my_rf1$importance %>% as_tibble() %>% bind_cols(as_tibble(colnames(new_data)[-6])) %>% arrang

```

7 References

- [1] J. C. Stoltzfus, "Logistic regression: A brief primer," *Academic Emergency Medicine*, vol. 18, no. 10, pp. 1099–1104, 2011.
- [2] T. Courville and B. Thompson, "Use of structure coefficients in published multiple regression articles: β is not enough," *Educational and Psychological Measurement*, vol. 61, no. 2, pp. 229–248, 2001.
- [3] C. H. Achen, *Interpreting and using regression*, vol. 29. Sage, 1982.
- [4] M. Mittlböck and M. Schemper, "Explained variation for logistic regression," *Statistics in medicine*, vol. 15, no. 19, pp. 1987–1997, 1996.
- [5] A. Agresti, *An introduction to categorical data analysis*. John Wiley & Sons, 2018, p. 109.
- [6] S. Menard, "Six approaches to calculating standardized logistic regression coefficients," *The American Statistician*, vol. 58, no. 3, pp. 218–223, 2004.
- [7] H. Midi, S. K. Sarkar, and S. Rana, "Collinearity diagnostics of binary logistic regression model," *Journal of Interdisciplinary Mathematics*, vol. 13, no. 3, pp. 253–267, 2010.

- [8] J. O. Rawlings, S. G. Pantula, and D. A. Dickey, *Applied regression analysis: A research tool*. Springer Science & Business Media, 2001, pp. 372–373.
- [9] T. P. Ryan, *Modern regression methods*, vol. 655. John Wiley & Sons, 2008, p. 170.
- [10] S. Sheather, *A modern approach to regression with r*. Springer Science & Business Media, 2009, p. 203.
- [11] T. T. Sajobi *et al.*, “Measures of relative importance for health-related quality of life,” *Quality of Life Research*, vol. 21, no. 1, pp. 1–11, 2012.
- [12] J. Bi, “A review of statistical methods for determination of relative importance of correlated predictors and identification of drivers of consumer liking,” *Journal of Sensory Studies*, vol. 27, no. 2, pp. 87–101, 2012.
- [13] J. W. Johnson and J. M. LeBreton, “History and use of relative importance indices in organizational research,” *Organizational research methods*, vol. 7, no. 3, pp. 238–257, 2004.
- [14] J. W. Pratt, “Dividing the indivisible: Using simple symmetry to partition variance explained,” in *Proceedings of the second international tampere conference in statistics, 1987*, 1987, pp. 245–260.
- [15] D. R. Thomas, E. Hughes, and B. D. Zumbo, “On variable importance in linear regression,” *Social Indicators Research*, vol. 45, no. 1, pp. 253–275, 1998.
- [16] A. E. Hoerl and R. W. Kennard, “Ridge regression: Biased estimation for nonorthogonal problems,” *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.
- [17] J. Bring, “A geometric approach to compare variables in a regression model,” *The American Statistician*, vol. 50, no. 1, pp. 57–62, 1996.
- [18] R. Azen and D. V. Budescu, “The dominance analysis approach for comparing predictors in multiple regression,” *Psychological methods*, vol. 8, no. 2, p. 129, 2003.
- [19] R. Azen and N. Traxel, “Using dominance analysis to determine predictor importance in logistic regression,” *Journal of Educational and Behavioral Statistics*, vol. 34, no. 3, pp. 319–347, 2009.
- [20] J. W. Johnson, “A heuristic method for estimating the relative weight of predictor variables in multiple regression,” *Multivariate behavioral research*, vol. 35, no. 1, pp. 1–19, 2000.
- [21] S. Tonidandel and J. M. LeBreton, “Determining the relative importance of predictors in logistic regression: An extension of relative weight analysis,” *Organizational Research Methods*, vol. 13, no. 4, pp. 767–781, 2010.
- [22] L. Breiman, “Random forests,” *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [23] D. R. Cutler *et al.*, “Random forests for classification in ecology,” *Ecology*, vol. 88, no. 11, pp. 2783–2792, 2007.
- [24] E.-H. A. E.-B. Rady, A. A. ELSheikh, and M. R. Ezzeldin, “MODIFIED DOMINANCE ANALYSIS.” *Pakistan Journal of Statistics*, vol. 36, no. 4, 2020.