

Table 1. Logistic regression analysis of estimators of having a stroke

| Variables | β^a | $\text{Exp}^b(\beta)$ | $\text{SE}^c\beta$ | P-value |
|--------------------------------------|-----------|-----------------------|--------------------|-----------|
| Intercept | -7.49 | 0.0005 | 0.358 | <0.001*** |
| Age | 0.069 | 1.071 | 0.005 | <0.001*** |
| Hypertension (vs. no hypertension) | 0.381 | 1.464 | 0.163 | 0.019* |
| Heart disease (vs. no heart disease) | 0.330 | 1.391 | 0.188 | 0.079 |
| Average glucose level | 0.004 | 1.004 | 0.001 | <0.001*** |

N = 5110 used in this logistic regression analysis; Residual deviance = 1591.5; AIC = 1601.5

^a: Coefficient of the estimator. ^b: Exponential value. ^c: Standard error.

*P-value < 0.05, ***P-value < 0.001

Table 2. Accuracy of the prediction of the testing data

| Types of Classification | Accuracy |
|--|----------|
| Logistic Regression ^a | 0.951 |
| Linear Discriminant Analysis ^a | 0.949 |
| Quadratic Discriminant Analysis ^a | 0.885 |
| K-nearest Neighbors Algorithm | 0.949 |
| Random Forest | 0.951 |
| Support Vector Machine | 0.951 |

^a: Using the same logistic model in Table 1