Hypothesis 1 HadHeartAttack ~ HighRisk + HighBMI + Diabetes

cross validation:

Call:

NULL

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) -3.14959 0.01335 -235.950 < 2e-16 \*\*\*

HighRisk -0.45607 0.05634 -8.095 5.73e-16 \*\*\*

HighBMI 0.06036 0.01844 3.274 0.00106 \*\*

Diabetes 1.28457 0.01973 65.120 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 104249 on 246021 degrees of freedom

Residual deviance: 100137 on 246018 degrees of freedom

AIC: 100145

Number of Fisher Scoring iterations: 6

**Diabetes** and **high BMI** are significant and **increase** heart attack risk. **High blood pressure** appears significant but shows a **negative** coefficient need to revisit its definition or checking interaction effects.

Logistic classification:

Confusion Matrix and Statistics

Reference

Prediction No Yes

No 46517 2687

Yes 0 0

Accuracy : 0.9454

95% CI : (0.9433, 0.9474)

No Information Rate : 0.9454

P-Value [Acc > NIR] : 0.5051

Kappa : 0

Mcnemar's Test P-Value : <2e-16

Sensitivity : 0.00000

Specificity : 1.00000

Pos Pred Value : NaN

Neg Pred Value : 0.94539

Prevalence : 0.05461

Detection Rate : 0.00000

Detection Prevalence : 0.00000

Balanced Accuracy : 0.50000

'Positive' Class : Yes

Linear Model Selection:

4 x 1 sparse Matrix of class "dgCMatrix"

s0

(Intercept) -3.13799389

HighRisk -0.36633099

HighBMI 0.03967641

Diabetes 1.27200346

The presence of Diabetes is the strongest indicator (coefficient = 1.27). HighBMI is positively associated. HighRisk surprisingly has a negative coefficient.