

MATH 4322 Final Project Group 9

Logistic Regression

```
library(readr)
cardio_train <- read_delim("cardio_train.csv",
  delim = ";", escape_double = FALSE, trim_ws = TRUE)
```

Rows: 70000 Columns: 13

-- Column specification -----

Delimiter: ";"

dbl (13): id, age, gender, height, weight, ap_hi, ap_lo, cholesterol, gluc, ...

i Use `spec()` to retrieve the full column specification for this data.

i Specify the column types or set `show_col_types = FALSE` to quiet this message.

```
heart.logistic = glm(cardio ~ ., family = "binomial", data = cardio_train)
```

Warning: glm.fit: algorithm did not converge

Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

```
summary(heart.logistic)
```

Call:

```
glm(formula = cardio ~ ., family = "binomial", data = cardio_train)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-8.4904	-0.9638	-0.0976	0.9896	4.6646

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-8.519e+00	2.148e-01	-39.669	< 2e-16 ***
id	1.685e-07	2.903e-07	0.581	0.562
age	1.488e-04	3.553e-06	41.887	< 2e-16 ***
gender	1.525e-02	2.107e-02	0.724	0.469
height	-5.729e-03	1.231e-03	-4.654	3.26e-06 ***
weight	1.535e-02	6.594e-04	23.275	< 2e-16 ***
ap_hi	3.953e-02	6.053e-04	65.314	< 2e-16 ***
ap_lo	3.001e-04	6.733e-05	4.457	8.33e-06 ***
cholesterol	5.233e-01	1.499e-02	34.913	< 2e-16 ***
gluc	-1.186e-01	1.700e-02	-6.978	3.00e-12 ***
smoke	-1.315e-01	3.317e-02	-3.965	7.33e-05 ***
alco	-1.691e-01	4.021e-02	-4.205	2.62e-05 ***
active	-2.098e-01	2.105e-02	-9.970	< 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: 97041 on 69999 degrees of freedom
 Residual deviance: 80920 on 69987 degrees of freedom
 AIC: 80946

Number of Fisher Scoring iterations: 25