Slajdy

Michał Maj

22 czerwca 2017

Slajd 1

Generalized linear model (GLM) is a flexible generalization of ordinary linear regression. GLM include:

- linear regression
- logistic regression
- Poisson regression
- gamma regression
- and more…

$$\begin{cases} y \sim ExpFamily(\mu(\theta), \phi) \\ g(\mu) = \eta = X\beta \end{cases}$$

Figure 1:

Slajd 2

Slajd 3

Slajd 4

Slajd 5

Slajd 6

GLM in H2O - Loss function + Elastic Net Penalty

For example in linear regression we have:

$$\min_{\beta,\beta_0} \frac{1}{2N} \sum_{i=1}^{N} (x_i^T\beta + \beta_0 - y_i)^T (x_i^T\beta + \beta_0 - y_i) + \lambda(\alpha\|\beta\|_1 + \frac{1-\alpha}{2})\|\beta\|_2^2$$

Slajd 7

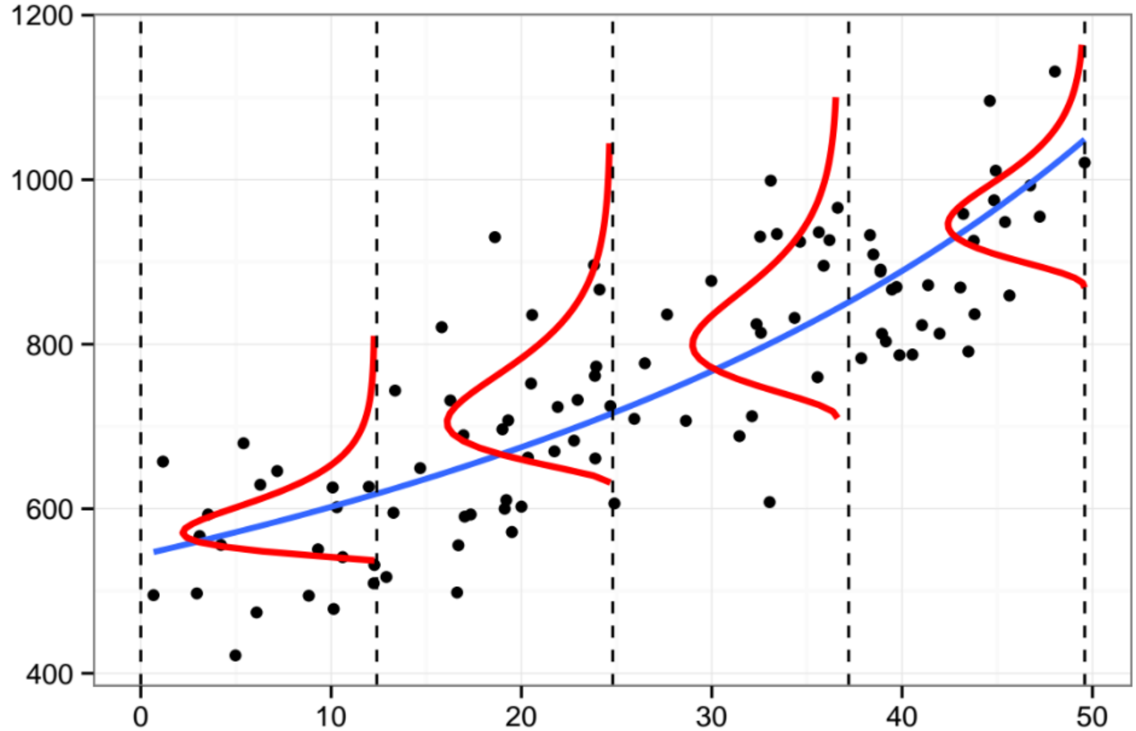

Figure 2: Visualization of GLM

$$\begin{cases} y \sim \mathcal{N}(\mu, \sigma^2) \\ \mu = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p \end{cases}$$

Figure 3:

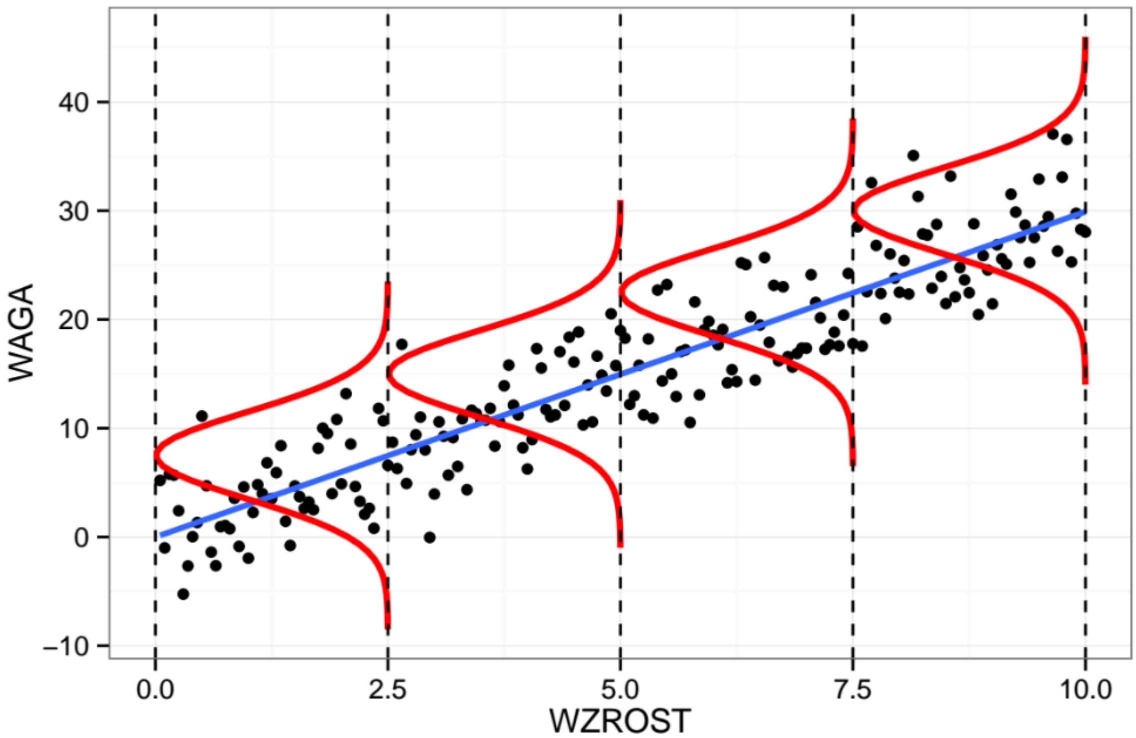

Figure 4: Linear regression as GLM

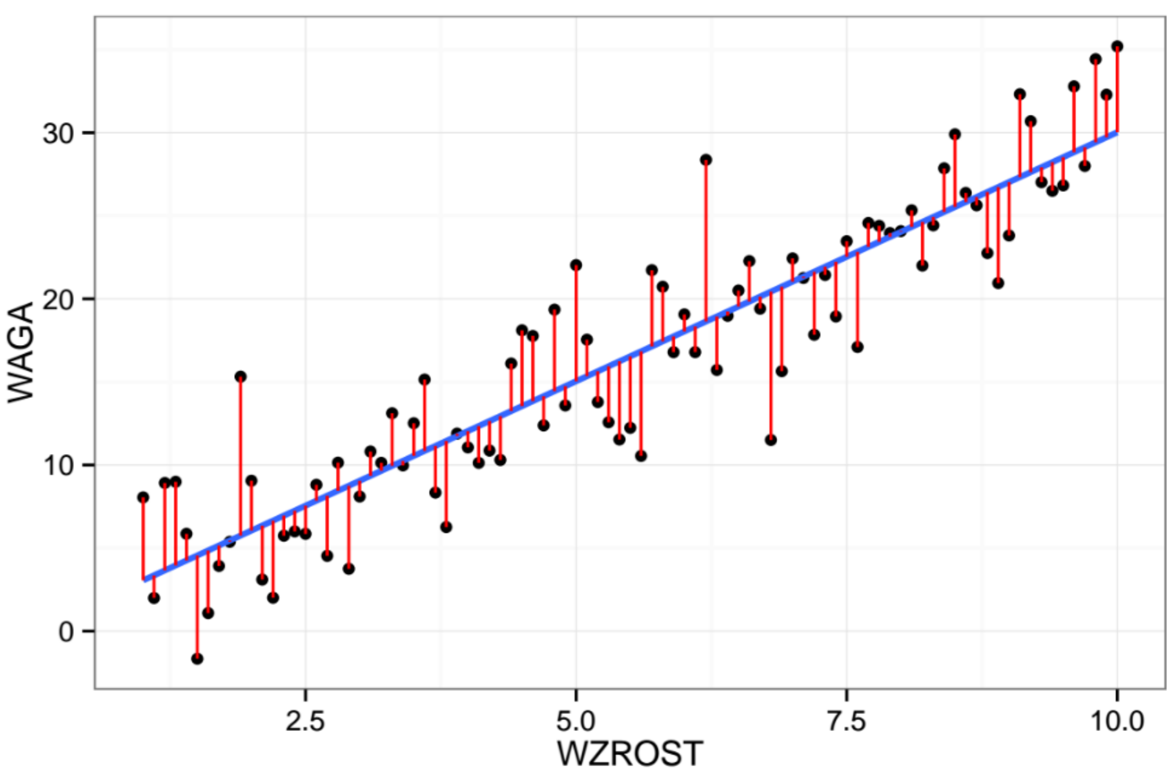

Figure 5: How to find Beta's ?

$$
\begin{cases} y \sim Bernoulli(p) \\ logit(\mu) = logit(p) = ln(\frac{p}{1-p}) = \eta = X\beta \end{cases}
$$

Figure 6:

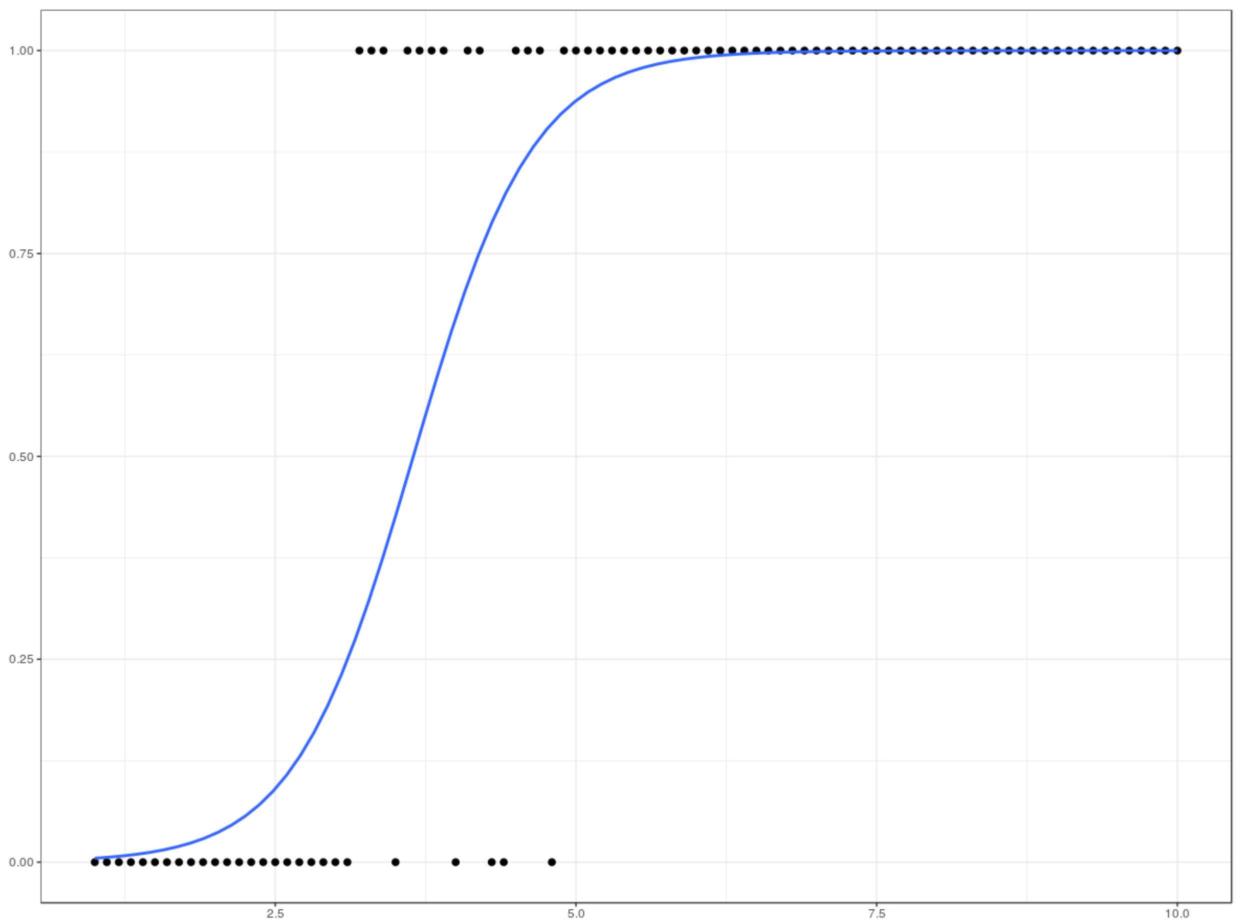

Figure 7: Logistic regression as GLM

$$
\min_{\beta,\beta_0} \frac{1}{N} ln(L(family, \beta, \beta_0)) + \lambda(\alpha\|\beta\|_1 + \frac{1-\alpha}{2}\|\beta\|_2^2)
$$

Figure 8:

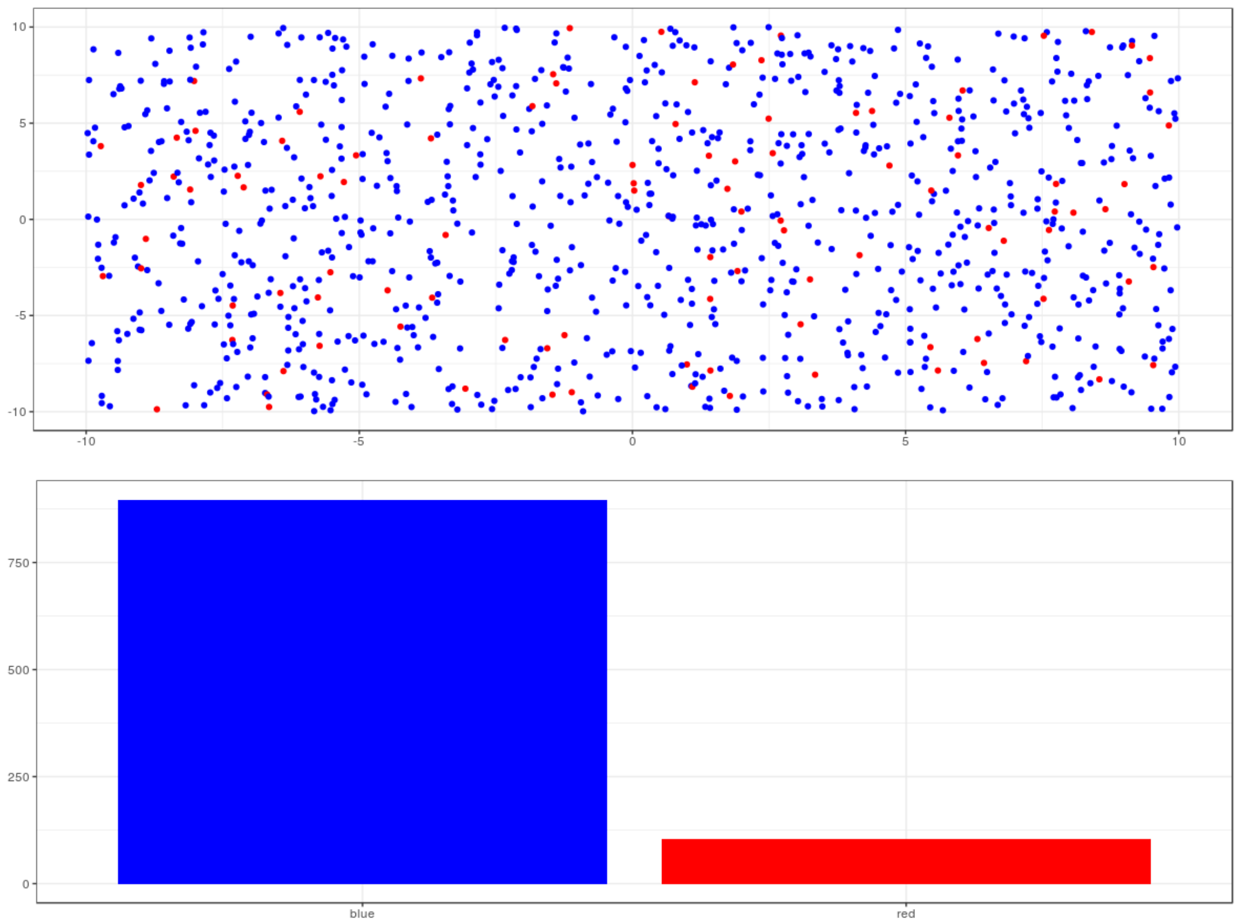

Figure 9: Orginal data

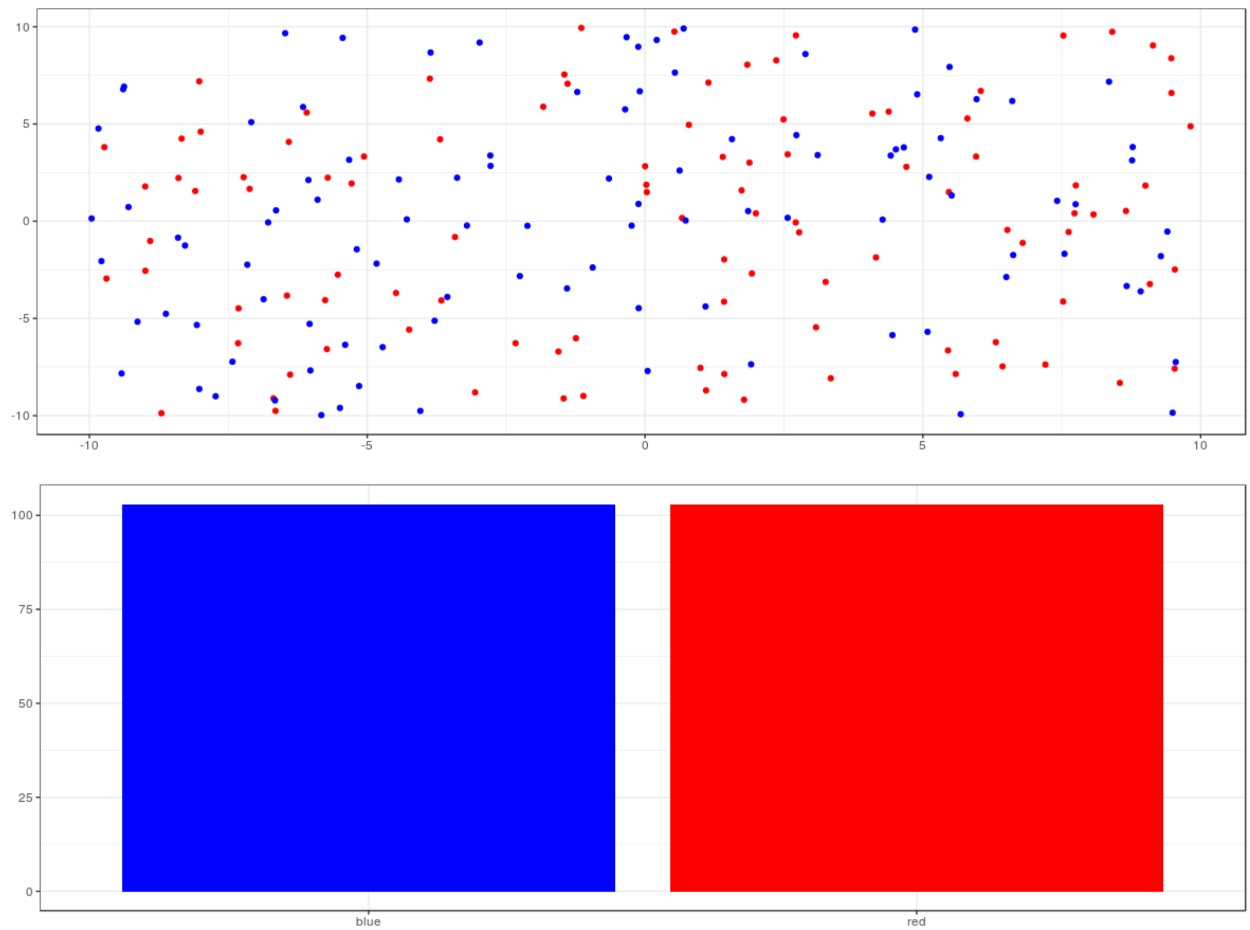

Figure 10: Undersampling

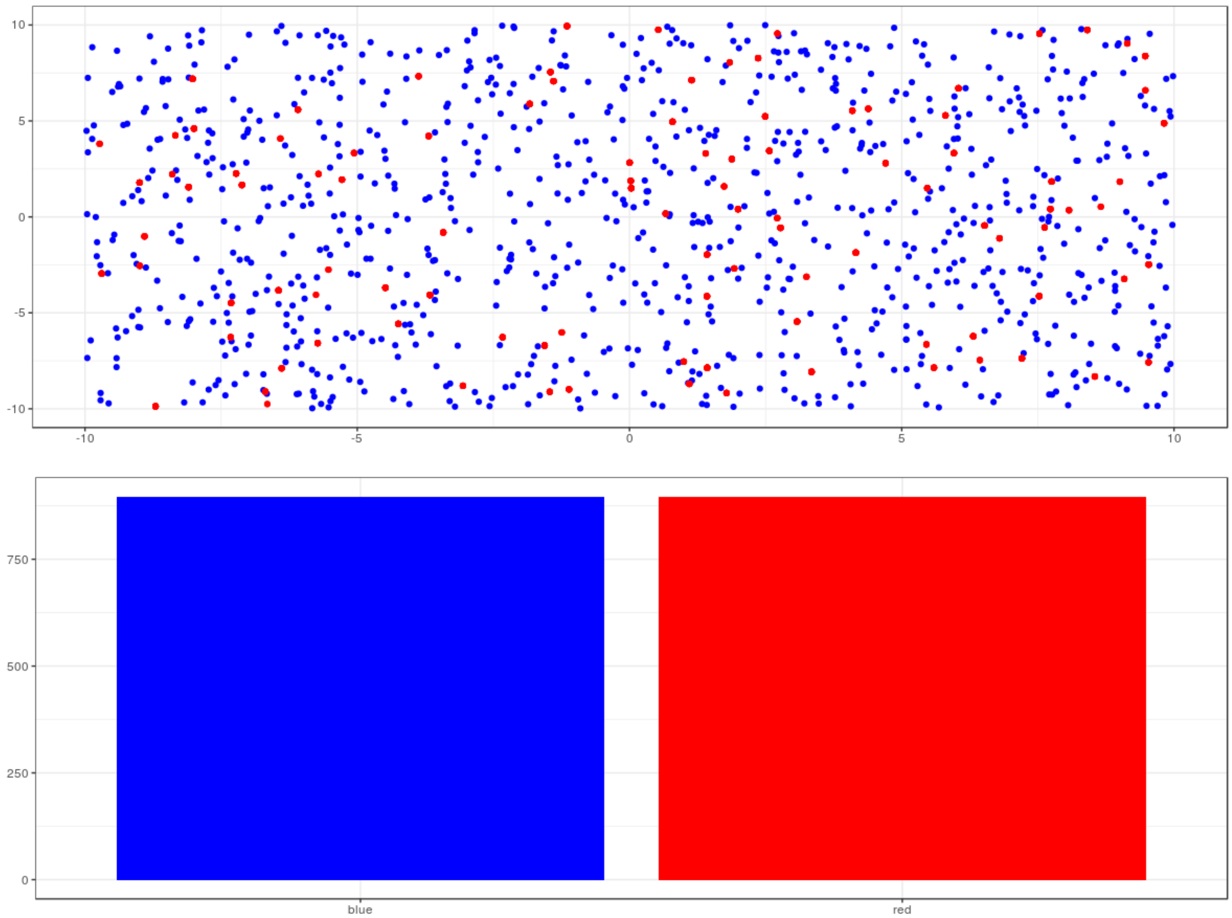

Figure 11: Undersampling

Slajd 8

Slajd 10

Slajd 11