glm is fitting generalized linear models. Glm always includes a random component, specifying the conditional distribution of the response variable given the predictors. When the response variable conforms to Gaussian distribution, the glm transfers to lm. In glm, the response variable depends on the predictors only through a linear function of the regression,. Glm is typically fit to data by the method of maximum likelihood, using the iteratively weighted least squares procedure. Denote the estimated value of the linear predictor as , because of the link function , the estimated mean of the response is and the variance of distribution is

gam is Fitting Generalized Additive Models: the function of this method is In our project, different features conform to different distributions, including both binomial and normal distribution, this method could build fit functions for different subjects and then add them together to get the final accurate predictions.

Glmnet is a generalized linear model fitting via penalized maximum likelihood, the regularization path is computed for the lasso or elasticnet penalty at a grid of values for the regularization parameter lambda.

To control the problem of over-fitting in the logistic regression, we could use different regularization methods by setting different alpha value. The penalty is defined as

.

We could also change the value of , in order to adjust different combinations of the norm and norm. Generally speaking, norm is more helpful when we intend to select useful features among all features. On the other hand, norm is more useful to solve the problem of over-fitting.

The super learner is the super learner prediction algorithm, which takes input X and Y and returns a predicted Y value. The above four algorithms in the SL.library have been chosen. The super learner will return the predicted value of Y depending on different algorithms and then compare the predicted value with the real value of Y aiming to determine the weighted of each algorithm. If in the training dataset, one algorithm has quite a bad performance, the weight of this algorithm in the final super learner function may approximate to zero. The loss function is defined as .